

## N O T I C E

THIS DOCUMENT HAS BEEN REPRODUCED FROM  
MICROFICHE. ALTHOUGH IT IS RECOGNIZED THAT  
CERTAIN PORTIONS ARE ILLEGIBLE, IT IS BEING RELEASED  
IN THE INTEREST OF MAKING AVAILABLE AS MUCH  
INFORMATION AS POSSIBLE

NASA CR 159682

MITRE-Bedford

MTR-3787

Vol. I

## FINAL REPORT

# Application of Advanced On-Board Processing Concepts to Future Satellite Communications Systems

(NASA-CR-159682) APPLICATION OF ADVANCED  
ON-BOARD PROCESSING CONCEPTS TO FUTURE  
SATELLITE COMMUNICATIONS SYSTEMS (Mitre  
Corp.) 406 p HC A18/XF A01

N80-12260

CSCL 17B

Unclass

G3/32 46239

J. L. Katz  
M. Hoffman  
S. L. Kota  
J. M. Ruddy  
B. E. White

JUNE 1979



MITRE

WRIGHT



MITRE Technical Report

MTR-3787

Vol. 1

## FINAL REPORT

# Application of Advanced On-Board Processing Concepts to Future Satellite Communications Systems

J. L. Katz  
M. Hoffman  
S. L. Kota  
J. M. Ruddy  
B. E. White

JUNE 1979

CONTRACT SPONSOR  
CONTRACT NO  
PROJECT NO  
DEPT.

NASA  
F19628-79-C-0001  
8680  
D-97

THE  
**MITRE**  
CORPORATION  
BEDFORD, MASSACHUSETTS

This document was prepared for authorized distribution.  
It has not been approved for public release.

## TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
LIST OF ILLUSTRATIONS	ix
LIST OF TABLES	xvi
SUMMARY	xxi
1 INTRODUCTION	1
1.1 TECHNOLOGY CONSIDERATIONS	2
1.2 STUDY OVERVIEW	4
2 GENERAL SYSTEM APPROACH	7
2.1 PROPOSED COMMUNICATIONS SERVICES	12
2.2 GROWTH CONSIDERATIONS	15
2.3 LIMITATIONS OF EXISTING SPACE ARCHITECTURES	18
2.4 ARCHITECTURAL CONSIDERATIONS	20
2.4.1 Architectural Options	22
2.4.2 Number of Options	26
3 SYSTEM IMPLEMENTATION ANALYSIS	31
3.1 SYSTEM CONTROL	31
3.1.1 Traffic	32
3.1.2 Terrestrial Signaling Schemes	32
3.1.3 Variable Assignment Techniques	36
3.1.4 Satellite Overlay Configuration	37
3.1.5 Access Schemes	41
3.1.6 Communication Protocols	42
3.1.7 Recommendations	56
3.2 MULTIPLE ACCESS/MULTIPLEXING	58
3.2.1 Overview of Techniques	58
3.2.2 Multiple Beams	66

## TABLE OF CONTENTS (Continued)

<u>Section</u>	<u>Page</u>
3.3 MODULATION/DEMODULATION	95
3.4 REGENERATION	102
3.4.1 Repeater Concept	102
3.4.2 Performance Analysis for Uncoded Data	105
3.4.3 Performance Analysis for Coded Data	112
3.4.4 Analysis of Results	124
4 BASEBAND PROCESSOR IMPLEMENTATION CONSIDERATIONS	137
4.1 ASSUMPTIONS	137
4.2 BASEBAND PROCESSOR IMPLEMENTATION AND TECHNOLOGY	140
4.2.1 Demodulation Processor Module	140
4.2.2 Digital Switch Interconnection	143
4.2.3 Downlink Modulators	144
4.2.4 Word Processor Architecture	144
4.3 MEMORY REQUIREMENTS AND TECHNOLOGY	145
4.4 LOGIC SPEED REQUIREMENTS	148
4.5 SUMMARY	152
5 MICROWAVE SWITCH TECHNOLOGY	159
5.1 RF SWITCHING ARCHITECTURE	159
5.1.1 Switch Technologies	165
5.1.2 Switch Architecture	168
5.1.3 Switch Architecture Realization	170
5.2 PIN DIODE TECHNOLOGY	202
5.2.1 Basic PIN Diode Operation	203
5.2.2 Single Pole-Single Throw Diode Switches	208
5.2.3 Multiple Diode Switches	214
5.2.4 Switching Speed and Biasing	217
5.2.5 Single Pole N Throw Switch	225
5.2.6 Driving Considerations	230
5.2.7 PIN Diode Cross Bar Conclusions	234

## TABLE OF CONTENTS (Continued)

<u>Section</u>	<u>Page</u>
5.3 FERRITE TECHNOLOGY	234
5.3.1 Microstrip and Stripline Circulators	236
5.3.2 Printed Circuit Circulators	239
5.3.3 Ferrite Switch Design	241
5.3.4 Ferrite Device Conclusions	243
5.4 OPTICAL COMMUNICATIONS TECHNOLOGY	244
5.4.1 Optical Switching	244
5.4.2 Switch Performance	250
5.4.3 Optical Switching Conclusions	254
5.5 MESFET SWITCHING TECHNOLOGY	254
5.5.1 GaAS MESFET Configuration	255
5.5.2 MESFET Operation	257
5.5.3 MESFET Switching Applications	263
5.5.4 MESFET Switching Conclusions	267
5.6 SWITCHING SUBSYSTEM CONSIDERATIONS	270
5.6.1 Broadcast Capability	270
5.6.2 Memory Control	271
5.6.3 Reliability	274
5.7 SWITCH STUDY CONCLUSIONS	279
5.8 RECOMMENDATIONS	280
6 COST PERFORMANCE TRADEOFFS	285
6.1 GENERAL CONSIDERATIONS	285
6.1.1 Implications of Power Flux Density Limit	287
6.1.2 Geostationary Arc Considerations	293
6.1.3 Link Imbalance Implications	295
6.1.4 Modulation and Signal Processing	297

## TABLE OF CONTENTS (Continued)

<u>Section</u>	<u>Page</u>
6.2 ACCESS METHODS	300
6.2.1 Case 1. TDMA	301
6.2.2 Case 2. FDMA	301
6.3 SATELLITE WEIGHT OPTIMIZATION	302
6.4 TERMINAL AND SATELLITE TRANSPONDER DESCRIPTION	308
6.4.1 Trunking Channel	309
6.4.2 Direct to User Channel	309
6.5 SATELLITE AND TERMINAL COST ESTIMATION	314
6.6 COST/PERFORMANCE TRADEOFFS FOR STRAWMAN SYSTEM	323
6.6.1 Case 1. Trunking plus Non-Regenerative TDMA Direct-to-User Transponder	326
6.6.2 Case 2. Trunking plus Regenerative TDMA Direct-to-User Transponder	328
6.6.3 Case 3. Trunking plus Regenerative FDMA Direct-to-User Transponder	330
6.7 TECHNOLOGY IMPLICATIONS	333
7 CONCLUSIONS AND RECOMMENDATIONS	335
7.1 MULTIPLE ACCESS/MULTIPLEXING	335
7.2 MODULATION/DEMODULATION	336
7.3 REGENERATION	339
7.4 BASEBAND PROCESSOR	339
7.5 MICROWAVE SWITCH	340
7.6 COST ANALYSIS	342

**"Page missing from available version"**

# LIST OF ILLUSTRATIONS (Continued)

<u>Figure</u>		<u>Page</u>
3-16	Processing Downlink Burst as Separated Uplink Sub-Bursts ( $B_u < B_d$ )	82
3-17	On-Board Processing Delay for Disparate Uplink and Downlink Burst Rates	84
3-18	Example Design Flow and Parameter Relationships	90
3-19	Derived Parameters for Design Example	92
3-20	Results of Sample Optimization	93
3-21	Results with Fewer Scanning Beams	94
3-22	Spectral Densities of SQPSK, MSK, and SFSK	97
3-23	Asymptotic Behavior of Mean-Square Crosstalk	98
3-24	Number of Regenerators per Beam for Signal Bandwidth $W = 300$ MHz and Carrier Packing Factor $\beta = 2/3$ ( $\beta W = 200$ Mbps)	101
3-25	Satellite Communications System	104
3-26	Probability of Error for QPSK	107
3-27	Performed Curves for Constant BER of $10^{-4}$ , HPA Backoff = 6 dB, TWTA Backoff = 4 dB	109
3-28	Performance Curves for Constant BER of $10^{-4}$ , HPA Backoff = 6 dB, TWTA Backoff = 0 dB	110
3-29	Performance Curve for Constant BER of $10^{-4}$ , HPA Backoff = 0 dB, TWTA Backoff = 0 dB	111
3-30	$R_o$ for Conventional (LRPT) Repeater with HPA Backoff = 6 dB, TWTA Backoff = 4 dB	117
3-31	$R_o$ for Conventional (LRPT) Repeater with HPA Backoff = 6 dB, TWTA Backoff = 0 dB	118

# LIST OF ILLUSTRATIONS (Continued)

<u>Figure</u>		<u>Page</u>
3-32	$R_0$ for Conventional (LRPT) Repeater with HPA Backoff = 0 dB, TWTA Backoff = 0 dB	119
3-33	$R_0$ for HDP with HPA Backoff = 6 dB, TWTA Backoff = 4 dB	120
3-34	$R_0$ for HDP with HPA Backoff = 6 dB, TWTA Backoff = 0 dB	121
3-35	$R_0$ for HDP with HPA Backoff = 0 dB, TWTA Backoff = 0 dB	122
3-36	$R_0$ for DEP with HPA Backoff = 6 dB, TWTA Backoff = 4 dB	125
3-37	$R_0$ for DEP with HPA Backoff = 6 dB, TWTA Backoff = 0 dB	126
3-38	$R_0$ for DEP with HPA Backoff = 0 dB, TWTA Backoff = 0 dB	127
3-39	Performance Curves for Constant $R_0$ of 1.4 Bits/Channel Symbol With HPA Backoff = 6 dB, TWTA Backoff = 4 dB	128
3-40	Performance Curves for Constant $R_0$ of 1.4 Bits/Channel Symbol With HPA Backoff = 6 dB, TWTA Backoff = 0 dB	129
3-41	Performance Curves for Constant $R_0$ of 1/4 Bits/Channel Symbol With HPA Backoff = 0 dB, TWTA Backoff = 0 dB	130
4-1	Processing Satellite System	138
4-2	FDMA/TDM Processor Configuration	139
4-3	Baseband Processing System	141
4-4	Memory Capacity Versus Access Time	147
4-5	Maximum Permissible Gate Delay as a Function of Clocking Rate	150
4-6	Speed-Power Product for Various Logic Families [Ranada, 1979]	151
4-7	Present and Projected Status of Subnanosecond Logic	153



## LIST OF ILLUSTRATIONS (Continued)

<u>Figure</u>	<u>Page</u>
5-1      Organizations Contacted	160
5-2      Interconnection of Fixed Geographical Regions	162
5-3      Frame Slot Organization	164
5-4      Candidate Media	166
5-5      Cross Bar Requirement	169
5-6      Fan-Out/Fan-In System	172
5-7      Single Pole Multiple Throw Switch	174
5-8      A 1 x 100 Switch Made of 1 x 10 Switches	175
5-9      Assemblage of SP100T SWs to Form a 100 x 100 Switch	176
5-10     Single Pole Multiple Throw Configuration	177
5-11     Performance of a 100 x 100 Cross Bar	179
5-12     Power Divider Configuration	181
5-13     Coupler Matrix External View	183
5-14     Microwaves News Release	184
5-15     Rearrangeable Switch Configuration	185
5-16     Physical Cross Bar and Feed Throughs	190
5-17     Coupler Cross Bar Realization	192
5-18     Cross Bar with Multiple Cross Points	194
5-19     Three Stage Connecting Network	198
5-20     Phased Array for TDMA Scan	201

## LIST OF ILLUSTRATIONS (Continued)

<u>Figure</u>	<u>Page</u>
5-21 Diode Switch	204
5-22 Simplified Equivalent Diode Circuit	204
5-23 PIN Diode Equivalent Circuit	206
5-24 Switching Time vs. Forward Bias and Peak Reverse Voltage	209
5-25 PIN Diode Switching Principles	210
5-26 Series Switch Performance	211
5-27 Shunt Diode Configuration	213
5-28 Shunt Diode Insertion Loss and Isolation	215
5-29 Shunt Diode Low Pass Matching	216
5-30 Shunt Diodes Quarter Wave Spaced	218
5-31 Series Diodes Quarter Wave Spaced	219
5-32 Series Shunt Switches	220
5-33 Shunt Diode Biasing and Switching	221
5-34 Reverse Recovery Time for Two Diodes	223
5-35 Clamping Susceptibility	224
5-36 Filtering Against Spikes	226
5-37 Single Pole N Throw Diode Switch	227
5-38 Grounded Wedge and Milled Decoupling	229
5-39 Sequence Digital Control, Decode, and Drivers	231
5-40 Driver for Shunt Switches HP	233

## LIST OF ILLUSTRATIONS (Continued)

<u>Figure</u>	<u>Page</u>
5-41 Circulation Principles	235
5-42 Y-Junction Microstrip Circulator	237
5-43 Stripline Y-Junction Circulator	238
5-44 Geometry of Center Conductor Pattern of Y Circulator	240
5-45 Y-Junction Ferrite Cross Point	242
5-46 Optical Switch Principles	245
5-47 Index Perturbation $\text{LiTaO}_3$ as Function of Voltage	248
5-48 Field Dependence of Critical Angles	249
5-49 Optical Cross Point	251
5-50 Wave Guiding and Switching by Internal Reflection	252
5-51 Optical Switch Electrode Pattern	253
5-52 MESFET Arrangement	256
5-53 Drain Current vs. Drain Voltage	258
5-54 MESFET Operation	259
5-55 MESFET Principles	261
5-56 FET SPST Configurations	264
5-57 Performance of Plessey MESFET as a Series Switch	265
5-58 Performance of 8 x 8 MESFET Switch	266
5-59 Dual Gate FET	268
5-60 MESFET Amplifier Performance	269
5-61 Wraparound for Broadcast	272

# LIST OF ILLUSTRATIONS (Concluded)

<u>Figure</u>		<u>Page</u>
5-62	RF Switch Control Memory	273
5-63	Wraparound Redundancy	276
6-1	Exemplar 20/30 GHz System Concept	286
6-2	Power Flux Density Limit Regions	288
6-3	Satellite Transponder Power and Antenna Diameter Required to Attain PFD Limit (Region 1) as a Function of Bandwidth at 20 GHz	290
6-4	Antenna Size vs. 30 dB Bandwidth	294
6-5a	Satellite Transmit Power vs. Trunking Terminal $G/T_s$	303
6-5b	Satellite Transmit Power vs. Direct-to-User Terminal $G/T_s$	304
6-6	Satellite Model	306
6-7	Trunking Transponder	310
6-8	Trunking Terminal	311
6-9	Non-Regenerative Direct-to-User Transponder	312
6-10	Direct-to-User Terminal	313
6-11	Regenerative TDMA Direct-to-User Transponder	315
6-12	Regenerative FDMA Direct-to-User Transponder	316
6-13	Antenna/LNA Cost vs. $G/T$	320
6-14	Transmitter Cost vs. Transmit Power	322

## LIST OF TABLES

<u>Table</u>	<u>Page</u>
2-1 Distribution of Standard Metropolitan Statistical Areas	10
2-2 Total Transponder Requirements for U.S. Domestic Services	17
3-1 Link Degradation Due to the HPA and Satellite TWTA Nonlinearity for BER = $10^{-4}$	105
3-2 Values of $E_b/N_0/R$ for Conventional Link Analysis at BER = $10^{-4}$	111
3-3 Uplink and Downlink Degradation Used in the Regenerative Repeater Case	112
3-4 Difference in Conventional and Regenerative Link $E_b/N_0$ Values for Uplink and Downlink	113
3-5 Difference in Conventional and Regenerative Link $E/N_0$ Values for Uplink and Downlink	131
4-1 On-Board Space Computer Systems	149
4-2 Comparison of Common Technologies Used in Implementing High Speed Arithmetic Devices	154
4-3 Comparison of Available Multiplier ICs	155
4-4 Comparison of Arithmetic Logic Units	156
4-5 Performance Comparison Between 8 x 8 and 16 x 16 Multiplication	157
5-1 Tentative Switch Specifications	171
5-2 Power Extracted for 100th Output Line	195
5-3 Diode Parameters and Performance	207
5-4 Parasitic Elements	207

# LIST OF TABLES (continued)

<u>Table</u>		<u>Page</u>
5-5	Equivalent Circuit Parameters of a Low-Noise GaAs MESFET	262
6-1	Computed System Noise Temperature for Indicated Link Availability	292
6-2	Minimum Size Antenna	293
6-3	Non-Regenerative Trunking Channel	298
6-4	Non-Regenerative Direct-to-User Channel	298
6-5	Regenerative Direct-to-User Channel, Hard Decision Processing	299
6-6	Regenerative Direct-to-User Channel, Full Demod/Remod	299
6-7	Direct-to-User Processor Comparison	300
6-8	Trunking Terminal G/T and G/T <sub>s</sub>	321
6-9	Direct-to-User G/T and G/T <sub>s</sub>	321
6-10	Relative Unit Costs	323
6-11	Non-Regenerative Trunking Transponder	324
6-12	Trunking Terminal Satellite and System Cost	325
6-13	Non-Regenerative Direct-to-User TDMA Transponder	326
6-14	Direct-to-User TDMA Terminal Costs, Case (1)	327
6-15	Direct-to-User System Costs, Case (1)	327
6-16	Minimum Cost System, Case (1)	328
6-17	Regenerative Direct-to-User TDMA Transponder	328
6-18	Direct-to-User TDMA Terminal Costs, Case (2)	329

LIST OF TABLES (continued)

<u>Table</u>		<u>Page</u>
6-19	Direct-to-User System Costs, Case (2)	329
6-20	Minimum Cost System, Case (2)	330
6-21	Regenerative Direct-to-User FDMA Transponder	331
6-22	Direct-to-User FDMA Terminal, Case (3)	331
6-23	Direct-to-User System Costs, Case (3)	332
6-24	Minimum Cost System, Case (3)	332

## SUMMARY

This study examines the effect of various forms of on-board processing on satellite communication system performance and cost in the 30/20 GHz band; appropriate techniques and critical technologies for the expected demand for domestic wideband trunking and direct-to-user services in the 1990's are recommended.

In order to ascertain and prioritize the technical requirements, a particular systems context and geostationary satellite architecture are assumed. High data rate common carrier traffic is concentrated by the terrestrial system into multiple 44 and 276 Mb/s digital trunks which may be radio frequency (RF) switched by satellite to connect approximately 40 major metropolitan centers of more than one million people. Lower rate carrier traffic characterizes direct-to-user services of the 1.5 and 6.3 Mb/s variety which may interconnect users directly via satellite signal regeneration, i.e., demodulation, baseband switching, and remodulation. High gain antennas capable of handling tens of simultaneous beams for efficient reuse of the 2.5 GHz wide frequency band, an RF switch with in the order of 100 input/output ports, and a digital baseband processor with gigahertz logic speeds and megabits of memory are viewed as key technological areas for satellite implementation.

Major results include: (1) detailed information on the state-of-the-art in RF switching technology and design recommendations for the on-board RF switch; (2) parametric relationships among system parameters which facilitate the determination of the number of fixed and scanning beams, the numbers of users by data rate, the relative allocation of system bandwidth to trunking and direct-to-user services, logic speeds, and memory sizes; and (3) preliminary system level cost/performance tradeoffs on the relationship between on-board signal processing and terminal size, complexity, and cost.

The principal conclusions advocate development of: (1) large satellite antennas with tens of fixed beams plus several scanning beams; (2) low power solid state 20 GHz GaAs field effect transistor (FET) power amplifiers in the 5 W range; (3) a 100 x 100 RF switch matrix for 500 to 600 MHz wide trunking channels; and (4) on-board transmultiplexing of frequency division multiple access (FDMA) uplink and time division multiplex (TDM) downlink direct-to-user channels consuming a few hundred megahertz of bandwidth and with data rates in the order of 15 Mb/s per beam. Commensurate RF trunking and direct-to-user terminal technologies are recommended. Together, (1) and (2) minimize satellite weight and imply lower terminal transmitter power. An estimated \$300M system cost savings is projected for (4) with FDMA and on-board baseband processing serving 5000 direct-to-user terminals.



## SECTION 1

### INTRODUCTION

Rapidly growing communications services are creating a demand for large capacity communications satellites which are efficient and economical. This demand, coupled with the need for cheaper terminals and the requirements of a mix of user types, means that future communications satellites must support a complex system structure. Technological development and implementation are clearly needed. In view of the increased importance of communications to national, economic, and social development, new initiatives in this area by the National Aeronautics and Space Administration (NASA) Lewis Research Center (LeRC) are necessary. To support the Application Division, this report examines the application of technology to on-board satellite processing, a technique which will strongly impact communications systems in the 1990's.

The focus of the study reported is on-board processing techniques and related technological issues regarding implementation of a multibeam communications satellite which would service a wide range of users in the 1990 time frame. The term on-board processing as defined in this study is limited in scope and is intended to refer only to the processing and/or switching of communications signals at intermediate frequency (IF) or baseband. That is, we are dealing only with the satellite subsystem between the receiver output and transmitter input. Although the processor types identified and discussed in this study are generally frequency band independent, we have emphasized techniques applicable to wideband multibeam communications systems operating in the 20/30 GHz band. The user segment of primary interest is characterized by a large number of terminals with different data requirements that access the satellite directly. Data rates of interest are 1.554 Mb/s to 274.176 Mb/s, representing standard digital T1 to T4 data rates utilized by common carriers in the United States.

NASA's interest in advanced system concepts for commercial communications satellites stems from the desire to identify the technology which needs to be developed in order to meet the demand for new services in the next two decades. Although it is clear from the literature that the need for communications services will grow, the specific type of service or user mix is a question which is being addressed by NASA/LeRC but which at this time is undefined. Thus, as a vehicle for identifying on-board processing techniques,

related system implications, and technology issues, an exemplar system was developed for analysis and discussion. This is only the first step in the evaluation of a system concept which will meet the need for new communications services in the 20/30 GHz band in an efficient and cost-effective way.

## 1.1 TECHNOLOGY CONSIDERATIONS

Communications services can be expected to make extensive use of both the 4/6 and 11/14 GHz bands in the 1980's with a particular emphasis on the higher bands for more specialized and tailored services, e.g., various forms of satellite data networks, where earth station size and frequency coordination problems are very important. It is also expected that as the demand for services expands, the 20/30/GHz band will emerge as the most appropriate band to provide these services; and that some operational 20/30 GHz systems, designed in whole or in part to provide data services, will emerge by 1990. Since the focus of this study is on-board processing, it is important to be cognizant of allied technological areas which need to be developed and which will impact the architecture of future systems. Developments in the technology areas of multiple access techniques, satellite antenna technology, satellite switching and processing, low-cost, highly reliable earth terminals, channel and source coding, and digital baseband processing equipment could be particularly significant in the development and evolution of systems with on-board processors. The following discussion highlights some of these areas and indicates the extent to which we consider each area.

In the area of multiple access techniques and modulation, some of the data services will be furthered by the introduction of the time division multiple access (TDMA) systems under development, such as the INTELSAT system. However, to provide all the services effectively will require a variety of access techniques, including various forms of TDMA, and bandwidth efficient modulation methods. Access techniques especially suited to the satellite networking and data collection services need to be studied, as does a very flexible and efficient demand-assigned TDMA or frequency division multiple access (FDMA) system. Some of these issues are addressed in later sections of this report.

If satellite communications experiences its expected overall growth, multiple frequency reuse will be required in many areas. Once the binary growth offered by polarization techniques is exhausted, spatial reuse techniques will be required and result in

multibeam satellites. Such satellites would be useful even in the absence of the frequency reuse issue as a means of concentrating the satellite energy where it is required and thus reducing earth station costs. This would be especially true in applications which potentially involve many earth stations. Although the satellite architecture considered in this report utilizes spatial reuse concepts, the technology associated with multibeam antennas is being addressed for NASA/LeRC by other contractors.

Use of multibeam satellites leads to a need for satellite switching and/or processing to provide the required interconnectivity among beams. This need is especially critical if the satellite plays a significant role in providing networking services or allocates resources of the satellite to various uplink communications requirements. These issues are the focus of this report and are discussed in detail.

The development of low-cost highly reliable earth stations, particularly in the 20/30 GHz band, could play a significant role in the development of service in this band. Note that the stations of interest here must, in general, have both transmit and receive capability. In particular, the development of very reliable transmitters, preferably solid state, is required. The earth station technology development and economics, including required maintenance costs, will play a major role in determining the feasibility of applications which involve a large number of earth stations. In this study, the impact of a processing satellite architecture on ground station costs is addressed.

Other elements of the overall system not addressed by this study but which will play an important role in determining whether economical communications services can be delivered are indicated here. Many of the digital services will be used between computers, at least potentially, with the result that very good error performance is required. Thus, the development of cost effective techniques and devices to provide error rates in the range of  $10^{-3}$  to  $10^{-12}$  is important. For example, this is the error rate range that appears to be a proper goal in an electronic mail system [Robbins, 1976]. For some applications, the development of very efficient source coding techniques to reduce the required transmission rate could have a large payoff. Particular examples of useful source coding include facsimile broadcast for remote publishing, the facsimile portion of an electronic mail system, and the video or facsimile transmission in a teleconferencing application.

These issues coupled with the data distribution requirement for high data rate systems seem to imply significant digital baseband processing capability at the earth stations. It is very important that the equipment be reasonably inexpensive and easy to maintain and yet able to do quite sophisticated processing. Modern solid state technology, if carefully applied, has the potential to satisfy most applications.

## 1.2 STUDY OVERVIEW

The work presented in this report is a first step in addressing the implications of on-board processing in future satellite communications systems. Many issues related to processing alone need to be studied, ranging from complex system questions to details associated with future device technology. Many of these issues cannot be specifically addressed at this time since only very general requirements for the system have been defined. However, based on a general system description, there are system and technology issues which can be examined. The approach taken in this study was to select for analysis specific issues related to system concepts, architecture, and technology which are important to the NASA/LeRC 20/30 GHz communications program. This report documents that study.

In section 2, the concept of an exemplar system is introduced to serve as the basis for discussing system issues. The environment in which the future system will operate is postulated along with a discussion of new communications services and expected growth in demand for services. Also included is a description of the method used to identify specific issues addressed in this report.

Section 3 considers system elements which impact on-board processor architecture. Control and access to the communications resource are discussed with major emphasis on a compatible overlay of the terrestrial network by the satellite system. An overview of multiple access and modulation techniques is presented first, followed by a discussion of applicable techniques for wideband application. A parametric analysis is carried out to gain an understanding of the relationship between system parameters such as numbers of antenna beams, burst rates, user numbers, throughput, etc. Performance analysis of a regenerative baseband processor is also presented.

Architectural issues related to implementing baseband processing are discussed in section 4. The implications of the analysis presented in section 3 as related to the baseband processor are addressed as a requirement. A discussion of projected logic and memory speeds as applied to the processor is also presented.

Section 5 describes radio frequency (RF) switching techniques and technology. A detailed analysis of switch architectures and components is presented and a unique concept for switch implementation is described.

A discussion of the relationship between on-board processing, and terminal size and complexity is presented in section 6. A technical discussion of satellite and terminal costs and cost tradeoffs of various technologies and system configurations is included. Important conclusions regarding system cost for processing and non-processing systems are summarized.

Conclusions and recommendations relating to each of the topics considered in the study are summarized in section 7.

Appendix A provides a detailed analysis of digital switching networks; appendix B examines multiple beam optimization; appendix C presents statistics on modulation characteristics; and appendix D discusses  $R_0$  as a channel performance measure.

References cited are listed at the end of each section. Volume II is an annotated bibliography of additional documents on the general subject of on-board processing techniques.

#### REFERENCE

Robbins, M., "Error Objectives of the Electronic Mail System,"  
Proceeding of IEEE International Conference on Communications, 14 -  
16 June 1976, pp. 15-7 to 15-9.

## SECTION 2

### GENERAL SYSTEM APPROACH

An exemplar candidate system configuration is chosen to focus the technical discussion in this report. A diagram of the exemplar system is shown in figure 2-1; it can be seen that the satellite is divided into two parts. The intent is to provide a system capable of handling a wide range of user data rates with specific parts dedicated to processing given types of data. Four distinct groups of users operating at standard transmit carrier data rates T1 (1.544 Mb/s), T2 (6.312 Mb/s), T3 (44.736 Mb/s) and T4 (276.176 Mb/s) are assumed.

Two generic classes of antenna for both the uplink and downlink segments on-board the satellite are also assumed. Fixed spot beam antennas would primarily be used in conjunction with the high data rate handling capability of the RF switch; a combination of fixed and scanning beams would be used with the demodulation/remodulation baseband processor. A judgment was made, based on analysis in this report, to limit the baseband processor to a data rate of T1 and T2 with the RF switch handling the higher T3 and T4 data rates.

Control of all satellite functions is concentrated in a central computer on-board the satellite. The functional operation of all peripheral processing units is controlled by the central computer. Primary control of the satellite is retained by a ground control facility which maintains communication with the on-board central computer through a command and control link.

The development of the exemplar system concept is based on several assumptions regarding the environment in which the system will operate. The ultimate system will be utilized by a purveyor of communications services such as a common carrier. The system will overlay an existing terrestrial network and must be compatible with it. Most important, the system must augment the capabilities of the terrestrial network by providing enhanced or new communications services in order to capture an economically viable share of the communications market. This last point is a key issue since the satellite system must compete with an in-place terrestrial network built with cheap dollars and essentially upgraded on a three-year basis to meet new demands. In addition, there are new improvements in digital microwave radios which will yield a tenfold increase in capacity, and developments in fiber optic cables which will make them very competitive with a copper cable system. Using new

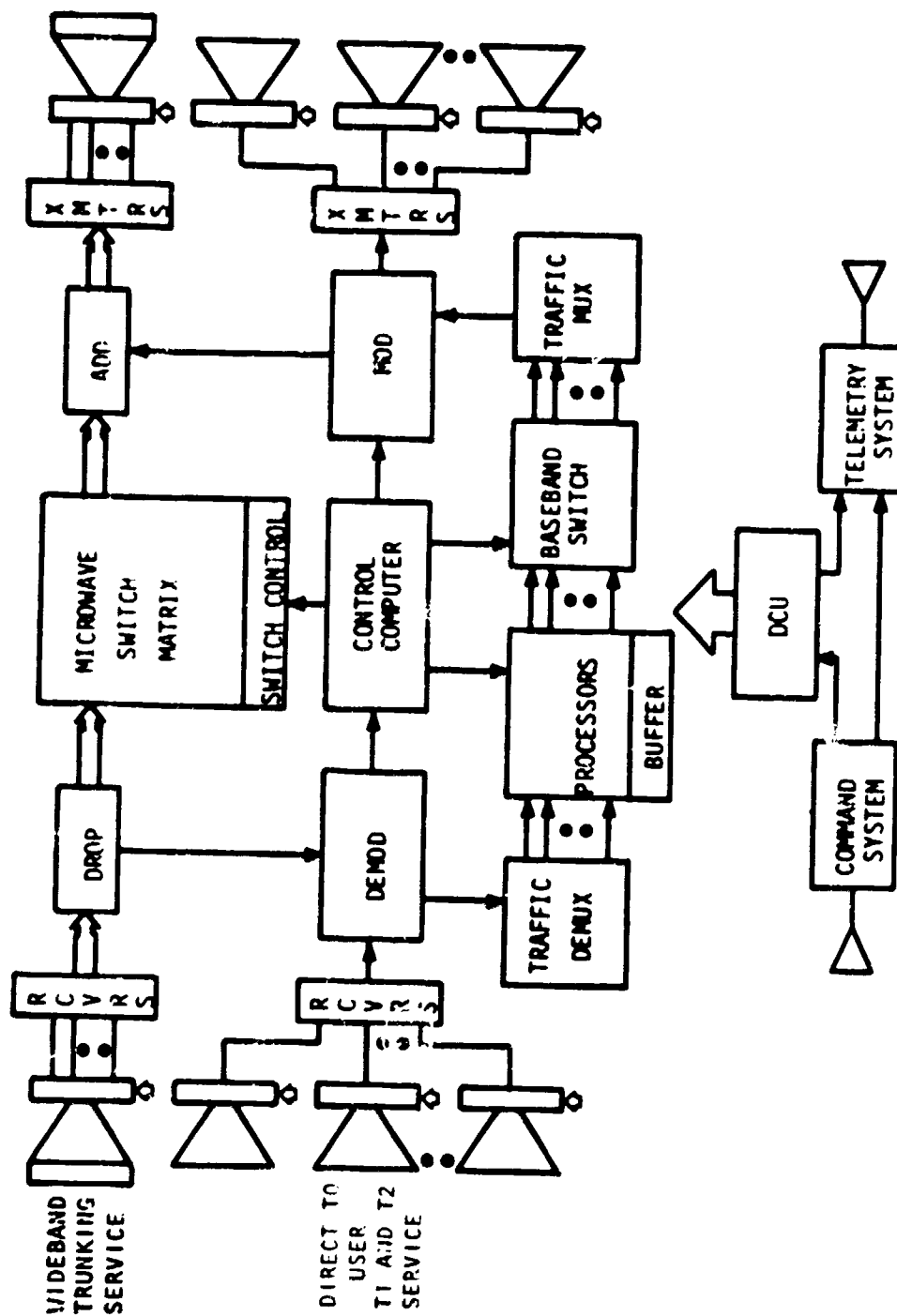


Figure 2-1. Processing Satellite System



satellite technology and on-board processing, advanced satellite systems will be an integral part of the communications network needed to meet the challenge for new and expanded communications services in the 1990's. For the purpose of establishing a contextual framework for this study, the remaining part of this section discusses communications services, growth, and technology which influence consideration of an advanced satellite system.

In the mid 1990's and beyond, processing satellite communications systems can be expected to utilize the following techniques:

1. Multiple up and downlink beams
2. Intensive frequency reuse
3. Flexibility in adapting to a non-uniform geographical traffic distribution

The design implications for each of these elements are strongly affected by the concentration of users in specific areas of the United States. A study of the 277 standard metropolitan statistical areas (SMSA) shows that these areas are scattered in a non-uniform pattern [Rand McNally, 1978]. There are 36 greater metropolitan areas in the United States with populations exceeding 1,000,000 and 37 metropolitan areas with populations between 500,000 and 1,000,000. The 277 SMSA in the United States are distributed as shown in table 2-1.

There is a well known city at the center of each SMSA which lends its name to the area. This city may serve as the concentrator for communications traffic arising in the area. It is reasonable to assume, in studying table 2-1, that fixed uplinks and downlinks will be assigned to 36 traffic generating areas whose populations exceed a million. The other 241 localities with lower traffic generating capacity will be served by a combination of fixed and scanning beams through a hierarchical system structure which concentrates data streams into a common wideband facility.

Table 2-1

Distribution of Standard Metropolitan Statistical Areas  
(Estimated 1 January 1978, Rand McNally)

<u>Population</u>	<u>Number of Areas</u>	<u>Sub Totals</u>	<u>Population</u>
Cities of Less than Half a Million			
0.5 - 1.0 x 10 <sup>5</sup>	24		
1.0 - 2.0 x 10 <sup>5</sup>	90		
2.0 - 3.0 x 10 <sup>5</sup>	46		
3.0 - 4.0 x 10 <sup>5</sup>	27		
4.0 - 5.0 x 10 <sup>5</sup>	17		
		204	43,900,000
Cities Between Half and One Million			
5.0 - 6.0 x 10 <sup>5</sup>	10		
6.0 - 7.0 x 10 <sup>5</sup>	9		
7.0 - 8.0 x 10 <sup>5</sup>	9		
8.0 - 9.0 x 10 <sup>5</sup>	7		
9.0 - 10.0 x 10 <sup>5</sup>	2		
		37	43,900,000
Cities with More than One Million			
1.0 - 2.0 x 10 <sup>6</sup>	21		
2.0 - 3.0 x 10 <sup>6</sup>	8		
3.0 - 4.0 x 10 <sup>6</sup>	2		
4.0 - 5.0 x 10 <sup>6</sup>	2		
5.0 - 6.0 x 10 <sup>6</sup>	0		
6.0 - 7.0 x 10 <sup>6</sup>	0		
7.0 - 8.0 x 10 <sup>6</sup>	2		
8.0 - 9.0 x 10 <sup>6</sup>	0		
9.0 - 10.0 x 10 <sup>6</sup>	1		
		36	92,000,000
TOTAL		277	161,850,000

The wideband multibeam satellite system is, therefore, assumed to operate in a network with a hierarchical system organization. The lowest level within the hierarchy is called the end user, and is viewed as a source or sink of information. The end user might, for example, employ a telephone, a video camera, a computer terminal, or a computer. End users are aggregated into user groups on the basis of common interests, geographical or political boundaries, or some other affinity characteristic. The user group is the lowest level of the network which will be able to access the satellite directly and is assumed to operate at a data rate of T1 (1.544 Mb/s). Data rates at T1 or T2 will be handled at the satellite by a baseband processor for maximum flexibility.

The next higher level of the hierarchy is called the local concentrator. Its exact nature depends on the service category and service under consideration, but it always plays the role of concentrating, interconnecting, or servicing user groups. Whether a user group accesses the satellite directly or is aggregated into a higher data rate by a local concentrator is a function of data requirements, economics, and terrestrial network topology. It is assumed that the local concentrator will access the satellite at either the T1 or T2 (1.544 or 6.312 Mb/s) data rate and be handled at the satellite by a baseband processor.

The next level of the hierarchy is called an area concentrator, and provides the function of interconnecting or providing service to a set of local concentrators. As with the local concentrator, the exact nature of the area concentrator is service category and service dependent. Data rates for the area concentrator will be at multiples of T3 or T4 (44.736, 274.176 Mb/s). Data rates at T3 or T4 will be handled at the satellite by an RF or IF switch matrix segment of the on-board processor and not demodulated to baseband.

As our interest is satellite communications, attention is focused on the use of a processing satellite to provide the required connectivity within the network. The connectivity specified in the following subsections is assumed to be achieved through the use of the satellite communications medium. Any other required connectivity is implicitly assumed to be achieved by terrestrial means.

## 2.1 PROPOSED COMMUNICATIONS SERVICES

Telephony, defined as the transmission of voice from point to point, has been the basis for most of the initial requirements in satellite communications. With the advent of new domestic systems, various other modes of connectivity are being considered beyond those required for a single voice channel. Our interest here is in dealing with data rates from T1 to T4 which correspond to 24 to 3600 voice channels.

In the U.S. there is a need for new services in two areas for satellite systems which handle T1 to T4 data rates. The first involves the provision of demand assigned trunk services to off-load heavily utilized long line circuits between switching centers. For example, to provide high quality voice grade service, the trunk line capacity between switching centers is configured so that no lines exceed 0.9 Erlang loading during peak usage hours. This means that there is a large excess capacity and costly physical plant not utilized in off hours whose existence adds to the cost of voice grade service. By having a demand assigned trunk line pool, available trunks can be assigned to switching centers as needed, thereby distributing the cost over a larger number of users. One of the reasons this type of service is a cost-effective application is the distance-independence of satellite communications costs. Existing terrestrial facilities have a per mile cost so that long distance trunks can be quite expensive. Thus, satellite links operating in the demand assignment mode present an economically competitive alternative to fixed terrestrial long distance trunks.

The second major growth area is the provision of domestic, private, or switched voice networks. Satellites appear to be uniquely suited to supplying economical intra-company communications for large dispersed corporations or communications between corporations comprising a common community of interest. This use of satellites is only beginning, but the market is potentially quite large.

Digital data and digital wideband services represent a large potential growth area for satellite communications. Many of the services are quite feasible today on a technical basis and should be economically viable under the right marketing arrangement. However, except for some low data rate and specialized high data rate services, most of the services are not widely available on satellite communications systems. Some of the applications are entirely new to satellite communications and will strongly involve the use of new technology, e.g., on-board processing, multiple access techniques, system organization, networking, and system control ideas.

This category of service is of specific interest to a 20/30 GHz system and is associated with a very heterogeneous set of applications ranging from computer networking to point-to-point video transmission. Some of the services in this category might well be integrated into a telephony based satellite communications system. As digital techniques become more prevalent in the transmission of telephony, the distinction between telephony and some forms of data communication, e.g., switched data service, will become increasingly blurred.

The list below characterizes services grouped in this category. The services range from the provision of digital pipes of varying rates to complicated complete networks. This list of services is only representative of the wide variety of services which belong to this class.

- Private Line - dedicated digital link connecting two user groups.
- Data Trunk - dedicated digital link connecting two local concentrators (switches) or area concentrators on a full-time (dedicated) basis.
- Switched Data Service - link actuated by data call. The link will typically be between two local concentrators, but it could be between two area concentrators. This service corresponds to demand access use of the satellite resource. The INTELSAT SPADE system for data is an example of this type of service.
- Data Collection - many-to-one connectivity with most primary data sources (users) being relatively low data rate and low duty cycle. The data sink is typically a centralized resource such as a large storage or data processing facility. Some topologies will concentrate a collection of data sources at an earth station. Dynamically shared use of the satellite channel is typical.
- Data Distribution - one to many broadcast of data. This service might be used to distribute data from a central location to users, perhaps after data collection and processing by a central collection point. Facsimile broadcast for remote publishing is another example application.

- Satellite Data Network - implies dynamically shared use of a common communications resource and use of the broadcast property or satellite switching to obtain connectivity. If the network is formed by use of point-to-point links connecting switches or concentrators, such use is considered under the data trunks or switched data service categories. The satellite network organization may range from fully distributed, with each user directly accessing the satellite, through varying levels of hierarchical structure with satellite access occurring after some terrestrial concentration and/or switching. The satellite data network service may be organized around a centrally located resource, such as a large computer facility, or the network resources may be distributed throughout the network.
- Teleconferencing - digital video transmission among n users. This may involve, for example, simultaneous availability of video signal from all participating sites at each site (fully connected network), transmissions from each site to a central site (conference director) with broadcast of a single transmission to all sites, or at any time a single site broadcasting to all other sites with the choice of the broadcasting site under the control of a central site (conference director). Other possibilities exist and the choice depends on the requirements of the specific application and many human factor issues, as well as economics.

The best organization of a satellite communications system, i.e., selection of services and options within services for supplying the data and/or wideband capabilities required to satisfy a specific application, depends on many factors. Some of these factors are: total traffic, traffic matrix, the requirements on response time, error rate and availability, the fundamental source characteristics in terms of data rate and duty cycle, the satellite and earth station costs associated with the various options for the current state of technology, and the availability and costs associated with terrestrial alternatives.

For many service applications discussed here, there is little experience on which to project demand. More definite organizational or service choices await appropriate detailed system tradeoff studies and a better understanding of the economics associated with satellite systems designed to provide particular services.

## 2.2 GROWTH CONSIDERATIONS

Growth in telephone service is easy to predict since there are good historical data on which to base projections. This is not the case for digital data and wideband services. With the large variety of services included in the data and wideband services category, it is reasonable to expect that there will be growth in the use of some of these services in the 1980's. On the other hand, it is very difficult to predict the total growth, or to decide with assurance which are the fastest growth services.

This uncertainty is due to at least three factors. First, the total market for data and wideband services has proved in the past to be difficult to predict accurately in the U.S. domestic environment. The newness and developmental state of some of the services, regardless of whether they are satisfied by terrestrial or satellite transmission, make market projections very subject to error. It was predicted in the 1960's that the volume of data services would exceed that of telephony services during the 1970's. Time has proved that prediction wrong, regardless of the measure of volume; similar predictions are now being heard applying to the 1980's. For example, it has been predicted that, for many large corporations, data communications expenditures will exceed expenditures for voice communications some time during the 1980's [Collyer, 1976].

A second factor that adds to the uncertainty of the satellite communications market for data and wideband services is that there is little history with respect to the provision of data services by satellite.

The third factor is the uncertainty of the market percentage that satellites will capture of those services where it must compete directly with terrestrial systems. This, of course, depends directly on the relative economics of the two competing techniques and on regulatory issues.

Terrestrial transmission technology can be expected to advance greatly in the 1980's with respect to conversion to digital techniques and increase in capacity. Promising technologies for the 1980's include single-sideband modulation techniques on microwave radio, optical fibers, and, perhaps, millimeter waveguides. Satellite systems have the advantage of quicker implementation and lower investment costs, particularly for geographically dispersed applications. They are also potentially very economically competitive with terrestrial systems for the provision of data services [Clarke, 1976].

Regardless of the uncertainty of hard predictions, it is clear that data communications will grow significantly in the 1980's. It is estimated that, at the present time, 15 to 20 percent of the INTELSAT traffic is data or message related, and this percentage can be expected to increase as new data services are offered.

An 18 to 20 percent compound growth rate until the mid-1980's is estimated for the U.S. domestic private line market with an estimated revenue of \$6 million by 1985 [Market Report, 1976]. Even as small a growth rate as 7 percent would still result in a \$2 billion business. It is probably reasonable to assume that approximately 20 percent of this market is data related.

Another estimate of data market growth is the growth in computer terminals and communications. This compound growth rate has been conservatively estimated at 15 percent per year, and other estimates are as high as 20 to 30 percent [Karp, 1976]. By 1985, terminals are expected to account for 25 percent of all computer equipment sales [Backler, 1976].

Table 2-2, which was taken from a Future System Inc., 1979 study, gives estimates of the total satellite 36 MHz transponder requirement for U.S. domestic service for the years 1978 to 1996. These requirements are based on the projection of both telephone and data traffic using a specific statistical model and assume that satellite systems which currently carry less than 2 percent of total long distance traffic will grow to only 8 percent in the year 2003.

While there is ample evidence that communications services will experience significant growth during the 1980's and 1990's, the real question is how much of this market satellite communications systems can capture.



Table 2-2

## Total Transponder Requirements for U.S. Domestic Services

<u>Mid Year</u>	<u>Number of Transponders</u>	<u>Annual Growth in Percent</u>
1978	41	
1979	70	70.7
1980	108	54.3
1981	146	35.2
1982	187	28.1
1983	224	19.8
1984	262	17.0
1985	298	13.7
1986	334	12.1
1987	368	10.2
1988	404	8.9
1989	433	7.2
1990	465	7.4
1991	498	7.1
1992	532	6.8
1993	576	8.3
1994	620	7.6
1995	665	7.0
1996	700	5.3

## 2.3 LIMITATIONS OF EXISTING SPACE ARCHITECTURES

The simplest kind of communications satellite is the repeater, or transponder, which amplifies and retransmits the received signal. This includes mixing the signal to a new carrier frequency to permit simultaneous reception and transmission. Three variations of the repeater concept are of interest: hard limiter, non-limiting amplifier with automatic gain control, and fixed-gain linear amplifier. Some of the inherent problems existing in conventional types of satellite repeater are indicated in this subsection. Discussion of the issue of connectivity in a multibeam satellite system is deferred. These problems provided the impetus for looking at on-board processing as a possible solution. Section 3 discusses methods for overcoming some of the problems through selection of appropriate multiple access method, modulation, and baseband processing. Qualitative assessments of each non-processing repeater type are included below.

The hard-limiting transponder has been the most popular, historically, for sound practical reasons: it is relatively easy to build and can be operated Class C or Class D for optimum d.c. to RF efficiency. A hard limiter is normally modeled as a device whose constant output power is divided among its output signals in a direct relation to the power ratios among its inputs. Specifically, the input and output proportions are nearly identical when the input signals are fairly similar in amplitude. When any one input signal is large compared to the sum of the others, its proportional share of the output power is increased by an additional factor. This signal suppression typically varies between 1 and 6 dB, depending on the character of the signals. Any interfering signal at the input also consumes a share of the output power. This effect is called power robbing, especially when the interferer is strong enough to subject the desired signals to small-signal suppression; it is one of the worst faults of the hard-limiting transponder. Because of this fault, an uplink interferer is doubly effective: beside masking the signal on the uplink, it degrades the signal-to-noise ratio on the downlink by reducing the relayed signal power below its clear-mode value. Even in a non-interfering situation, power imbalance can be a serious problem; a strong signal in an FDMA net of simultaneous users monopolizes the output power, to the detriment of the small users. This can only be remedied (in the clear mode) by power control, in which each user's transmitted power is adjusted to maintain predetermined ratios among the signals arriving at the satellite. A user's uplink power is adjusted equivalently to keep his share of the downlink power proportional to his data rate.

The other major fault of the hard-limiting transponder is generation of intermodulation (IM) products among signals of differing carrier frequencies. For a very small number of simultaneous users, the carrier frequencies can usually be selected so that IM products of all significant orders are outside the passbands of the user signals. As the number of users increases, the frequency-assignment problem becomes very difficult; some IM products fall in the downlink passbands, where they rob power and degrade signal-to-noise ratios. Worse still, they can fall in the receive band and degrade the uplink signals.

The second type of transponder is the non-limiting amplifier with automatic gain control (AGC). If it could be realized perfectly, such a device would not have the small-signal suppression and intermodulation distortion that result from saturation or clipping in a limiter. The associated disadvantage is that although the final power amplifier would be operated at a constant output level, in practice it must be backed off several decibels from saturation in order to operate Class A. It would therefore have lower capacity and lower d.c. to RF efficiency than a hard limiter. Another disadvantage is that, in spite of its linearity, it still requires power balance on the uplink signals. Also, there would be serious problems in practical realization of a linear/AGC transponder with the necessary wide dynamic range and good linearity in both the amplifier and the AGC loop.

A linear fixed-gain transponder design would offer two major advantages: it is inherently very simple, and it essentially avoids power balance problems. However, the latter property is achieved by setting its gain so that the total RF output power under normal uplink signal conditions is significantly below the maximum available. Thus, any signal in the output can increase without decreasing the others, provided that the nominal RF power level is low enough to accommodate the increase without saturation. On the other hand, to obtain adequate capacity and efficiency, the nominal output power level should be as high as possible. These conflicting requirements can probably best be reconciled in practice by carefully calibrating system gains and losses, and by exercising a moderate degree of power balance.

## 2.4 ARCHITECTURAL CONSIDERATIONS

This study looks at a broad range of processing satellite concepts for meeting system needs, as well as problems in the current generation of communications satellites. Many possible architectures must be evaluated when considering future systems and future technology capabilities. At this point, only limited system requirements have emerged to serve as the basis for making a decision on specific system alternatives. This subsection enumerates some of the satellite processing possibilities, provides a rationale for emphasizing one approach over another, and justifies focusing this study on specifically selected issues instead of attempting an exhaustive analysis of all architectural candidates on a global scale.

Although the vehicle for discussion is a tree structure not unlike that considered by other authors, e.g., [Heggstad, 1976], the specific options are tailored to the problem at hand, namely, the realization of wideband domestic telecommunications services via a satellite platform over the U.S. The tree grows from a base of trunked user services, through a series of earth terminal, uplink, processing satellite, and downlink branches. Each path through the tree defines a distinct architectural configuration.

A truncated version of this tree is shown in figure 2-2. Refer to the legend; a lower bound on the number of branches extending immediately above a fork in the tree is indicated beside the associated branchpoint. For now, the other numbers and circles around numbers should be ignored. They are explained later when the number of choices available is computed.

Each branch is labeled by a boxed caption descriptive of an architectural option. Branch arrows indicate options that should be considered in the present study; branches without arrows indicate topics which are outside the scope of this report. Heavier shaded arrowheads point along favored paths through the tree which receive greater attention than other paths; only one branch is favored at any branchpoint. The decision to favor a particular branch is generally dependent upon a preceding subsequence of options. For example, TDMA is favored for a fixed beam uplink and FDMA is favored for a scanning beam uplink. Less obviously, quadri-phase shift keyed (QPSK) uplink modulation is preferable to minimum shift keyed (MSK) with TDMA to a demodulating satellite. The reasons for such choices are covered in greater detail below and in subsequent paragraphs.

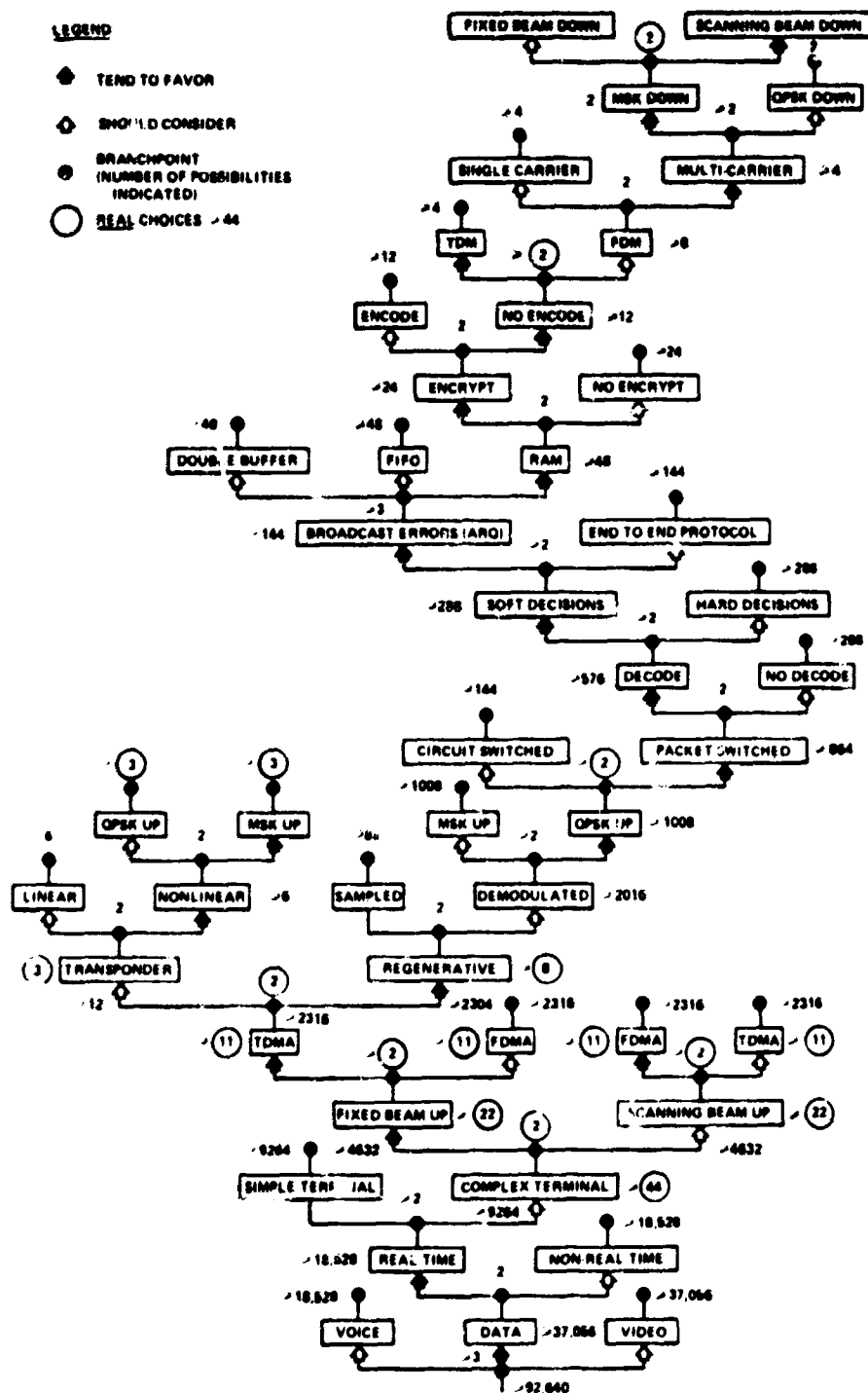


Figure 2-2. A Tree Structure of Satellite Processing Architectural Alternatives

The tree is truncated at certain branchpoints to avoid having to duplicate superstructure appearing above another branch on the same tree level. Only a relatively few paths of interest are shown for the sake of illustration.

The tree is ascended while some of the architectural options and preferences are explained. Then the number of choices is computed as the tree is descended.

#### 2.4.1 Architectural Options

Begin at the tree base; three kinds of user service are considered: voice, data, and video. Digital data service is favored for analysis because it comprises elements of both digital voice and digital video services. Continue up the data branch; real time data are preferred in order to retain the full impact of the 1/4 sec round trip propagation delay inherent in geostationary satellite communications. However, non-real time applications such as facsimile should also be considered.

Narrowband services are not considered in this study. Small mobile earth platforms, for example, requiring simple terminals, are outside the scope of this report.

Fixed beam uplinks are preferred to scanning beam uplinks; a higher useful data rate is possible if a given terminal is constantly illuminated by a beam, all other considerations being equivalent. On the other hand, scanning beams may be one method of reducing satellite complexity or accommodating varying traffic loads in fixed beam regions or remote areas. The circle around the "2" at this branchpoint of figure 2-2 means that both the fixed and scanning beam options should be examined with special care; there is an important choice to be made here.

At the next tree level, one may consider at least two generic accessing schemes: TDMA and FDMA. Again, these represent important options that must be considered in some depth. TDMA is favored for terminals in a fixed uplink beam and FDMA is favored for those in a scanning uplink beam because these choices tend to equalize the burst rate demands on the terminals in the two kinds of uplink beam. However, since this rationale may be critical only when unattainable burst rates are otherwise implied, other combinations of multiple accessing and uplink beams should also be considered.

In the context of this report, there are only two distinct families of processing satellites: transponders and regenerative

repeaters. Both families must be treated: transponders primarily for the RF or IF switching applications, and regenerative repeaters for the demodulation/remodulation features. Regenerative repeaters are given preferential treatment because of the larger number of processing options available and the greater potential for enhanced performance for lower data rate users.

Follow the transponder tree branch for the moment. There are two basic possibilities for the repeater: linear and nonlinear. The nonlinear option tends to be favored in practice in order to achieve more output power. Efficient linear amplification is a well known area for research and development that would alleviate many of the adverse effects of nonlinear transponders. For example, FDMA could be used without power control at the transmitters, and intermodulation problems on the downlink side of the satellite would be minimized.

Various uplink modulation schemes can be considered at the next tree level. If TDMA is used with a nonlinear transponder, a constant-envelope modulation scheme such as MSK, offset QPSK, or even a single carrier frequency division multiplex (FDM) signal may be preferable to a non-constant envelope modulation to minimize amplitude modulation (AM) to phase modulation (PM) distortion effects in traversing the nonlinearity of a traveling wave tube (TWT), for example. In the case of a bandlimited channel with an ideal bandpass limiter characteristic of solid state amplifiers, MSK or offset QPSK seems superior to QPSK [White, 1978] but the peculiarities of a soft limiter may favor QPSK. The modulation question needs further examination after the filtering and nonlinear characteristics of the transponder channel are established.

The only options to be considered above the uplink modulation level in the transponder subtree are the RF (or IF) switching alternatives. The modulation format cannot change, and it is assumed that only fixed downlink beams apply for the wideband RF switched service. Several switching alternatives are discussed in section 5.

Return to the regenerative tree branch; only two options are recognized: demodulation and sampled data. The latter refers to the possibility of downlink transmissions that are digital but not fully demodulated versions of the uplinks [Sarles, et al., 1977]. Although this may be simpler than a satellite group demodulator of several uplink signals, additional downlink bandwidth is generally required. Such a sampling regenerator is not considered in this report and is mentioned only for the sake of completeness.

When a multiple access decision has been made for the demodulated subtree, the range and/or sensitivity of uplink modulation choices can be narrowed considerably. For example, with TDMA and a regenerative repeater, the system designer might be relatively indifferent to the uplink modulation choice; any bandwidth and detection efficient modulation such as QPSK is probably satisfactory. Since QPSK is simpler to implement than MSK, QPSK is preferred in this case. On the other hand, if FDMA were selected, MSK might be preferred because of its superior crosstalk properties when the uplinks are at different burst rates or uncoordinated in network symbol timings, frequency offsets, and transmitter powers. MSK is more complex and therefore more difficult to implement than QPSK. Other modulation schemes are covered in section 3.

At the next higher tree level, one could decide the type of on-board switching employed. It is assumed that there will be no decoding or encoding in the satellite if circuit switching is selected. On-board coding, at least on a single channel basis, may not enhance performance sufficiently to be worth the increase in satellite complexity. If coding is necessary, end-to-end coding should be adequate with circuit switching.

In packet switching, an individual data stream is parsed into convenient block sizes at the transmitter. This presents the opportunity for link-by-link coding since at least the header portion of a packet must be recognized at each receiving network node. Here satellite decoding permits the acknowledgment, reformatting, and routing of packets based on information content which can lead to more efficient use of the satellite channel. For instance, an automatic acknowledgment scheme might involve the downlink retransmission of only a reformatted header to the user sending the corresponding uplink packet. At least two families of packet protocols, e.g., satellite broadcast or end-to-end, could be applied. Thus, in the packet switched subtree, a coding option is allowed. Soft decision decoding is often at least 2 dB superior to hard decision decoding in terms of the required signal-to-noise ratio at the receiver.

The on-board coding question could also be considered before the form of on-board switching. In the context of this study, it seems more natural to discuss switching first.

To take advantage of the fundamental broadcast nature of the satellite medium, the automatic repeat request (ARQ) option of the downlink broadcasting of information specifying uplink packets in error is followed. Next, there are at least three ways of handling



on-board storage requirements: double buffer schemes, where the last packet is transmitted on the downlink while the next uplink packet is received, first-in first-out (FIFO) buffers using shift registers, and random access memory (RAM). These alternatives are listed roughly in order of increasing memory size and complexity. A random access memory organization is tentatively favored because of the increase in flexibility afforded with on-board microprocessor implementations.

Downlink encryption is not advocated; it is mentioned as a possible option in view of current industry concerns in the area of public key cryptography [Hellman, 1978]. Since encryption is relatively easy to implement, it could be provided by the satellite as a downlink service. Individual users could be satisfied with this downlink cover or, in addition, provide their own end-to-end encryption.

At the next level of the tree, a decision is made on whether to use downlink coding. The added complexity is probably not worth a minimal improvement in performance, even with packet switching, since different packets generally collide only on the uplinks.\*

There are at least two options for downlink multiplexing, time division multiplex (TDM) and FDM, with FDM further divisible into single carrier or multi-carrier forms. The uplink preference is TDMA; TDM is preferred on the downlink to help equalize the uplink and downlink burst rates. Similarly, FDM would be used with FDMA under this criterion. Nevertheless, both TDM and FDM should be seriously considered, regardless of the accessing scheme, since other criteria may dominate the desire to equalize burst rates. For example, multi-carrier FDM may imply simpler downlink receivers while TDM may simplify on-board intermodulation problems.

A single carrier FDM downlink would avoid on-board intermodulation products but might introduce crosstalk from adjacent sidebands in the downlink receivers. Multi-carrier FDM would relieve any crosstalk problem on the ground but may aggravate the spacecraft intermodulation problem, depending on the number and frequencies of the downlink carriers. The multi-carrier option is favored here in the interest of terminal simplicity and performance.

---

\* The form of packet switching accepted as an option in this report precludes packet collisions on the uplinks by virtue of the TDM nature of each packetized uplink data stream. Thus, the need for uplink packet acknowledgment and reliability coding is reduced by the absence of this contention factor.

The downlink modulation is largely determined by the multiplexing scheme selected, just as the uplink modulation was affected by the multiple access choice. With multi-carrier FDM, MSK would be significantly better than QPSK because of the superior crosstalk (interchannel interference) properties of MSK. This presumes that the downlink carriers are not all modulated in synchronism, e.g., different burst rates can be employed. Note that if the downlink signals were perfectly coordinated in symbol timing, frequency uncertainty, and transmitter power, the downlinks could be made orthogonal and the modulation choice would be less important.

Finally, at the top of the tree there is the option of fixed or scanning downlink beams. A scanning beam preference is indicated for an FDM downlink just as a scanning beam uplink was favored with FDMA. Again, the "2" at the branchpoint is circled to indicate that both options should be seriously considered.

#### 2.4.2 Number of Options

The method of computing the lower bounds on the number of options shown next to terminated branchpoints or beside various other branches of the subtree of figure 2-2 is explained below. This computation requires a traversal from top to bottom.

Begin at the top level of the subtree; the fixed and scanning downlink beam option accounts for two possibilities at the MSK down branch and the QPSK down branchpoint. Since more than two downlink modulations are possible, there are at least four options at the multi-carrier branch, a lower bound obtained by summing the number of options associated with the adjacent ascending branches. The identical subtree can be visualized above the single carrier branch, so at least eight combinations exist above the FDM branch.

It is assumed that the single carrier alternative for TDM is the only viable downlink option. This suggests a lower bound of only four at the TDM branchpoint, and 12 at the no encode branch. An identical subtree would exist above the encode branch, which implies a lower bound of 24 at the encrypt branch.

Proceeding downward in a similar fashion, one obtains a lower bound of 576 at the decode branch. Only 288 is indicated at the no decode branchpoint because hard decisions are assumed. The lower bound at the circuit switched branchpoint is only 144 because no coding or protocol options are considered. Descending further, one

finds that the lower bound of 288 at the sampled branchpoint arises from a subtree identical to that along the circuit switched path.

Note that the transponder superstructure contains many fewer options than the regenerative option, as mentioned earlier. Finally, the voice branchpoint has half the lower bound as data or video branches since the real time alternative is considered mandatory with voice.

Thus, there are at least 92,640 distinct architectural configurations or paths in the tree of figure 2-2. Of these, there are really only 44 options that might be considered seriously. The smaller number was computed by descending the tree in the same fashion but counting as true options only the alternatives associated with branchpoints having a circled lower bound.

In particular, above the regenerative branch, the viable alternatives include fixed or scanning downlink beams, TDM or FDM, and circuit or packet switched, for a subtotal of at least eight possibilities. The transponder path yields at least three RF switching combinations of import. Hence, there are at least 11 viable options for each multiple access scheme. Since every multiple access scheme must be considered, as well as both fixed and scanning beam uplinks, one obtains the true lower bound of 44.

Even 44 architectural alternatives are too many options to be adequately treated on an individual basis in this report. Instead, the most viable alternatives have been selected and the relative merits of fixed or scanning beams, FDMA or TDMA, circuit or packet switching, and QPSK or MSK, etc., are considered in section 3 in conjunction with RF switching and baseband regeneration.

## REFERENCES

- Backler, J., "Terminals--Net Growth Spurs Communications Satellites," Minicomputer News, 25 March 1976.
- Clarke, A. C., "Preparing for the Data Deluge," Data Communications, May/June 1976, pp. 11-12.
- Collyer, E. A., Speech before the American Institute of Industrial Engineers (AIIE), Federal Government Data Systems Conference, 7 April 1976.
- Future Systems Inc., "Large Communications Platforms Versus Smaller Satellites," February 1979.
- Heggestad, H. M., "Analysis of UHF MILSATCOM Architectural Alternatives, Volume I: General System Considerations," TN 1976-41, Vol. I, Lincoln Laboratory, M.I.T., 26 October 1976.
- Hellman, M. E., "An Overview of Public Key Cryptography," IEEE Communications Society Magazine, Vol. 16, November 1978, pp. 24-32.
- Karp, H. R., "System Network Architecture--Building Blocks to Support Future Business," Data Communications, January/February 1976, p. 19.
- "Market Report: Satellite Business Systems," Telecommunications, June 1976, p. 16.
- Rand McNally Map of Standard Metropolitan Statistical Areas 1978.
- Robbins, M., "Error Objectives of the Electronic Mail System," Proceedings of IEEE International Conference on Communications, 14-16 June 1976, pp. 15-7 to 15-9.
- Sarles, F. W., Jr., L. W. Bowles, L. P. Farnsworth and P. Waldron, "The Lincoln Experimental Satellites LES-8 and -9," EASCON-77 Record, 26-28 September 1977, Arlington, VA, pp. 21-1A through 21-1U.
- Sedlacek, W. C., Leonard, R. E., Burtt, J. E., "Summary Report: Information Transfer Systems Requirement Study," NASA CR-73425, March 1970.

Staelin, D. H., "Expanding Broadband Switched Communications Networks," Satellite Communications, Vol. 3, January 1979, pp. 26-30.

Yeh, Y. S., Reudink, D. O., "The Organization and Synchronization of a Switched Spot-Beam System" ICCS Proceedings 1978.

White, B. E., "Simulation of Two Bandwidth Efficient Modulation Efforts in Satellite Communications," Proc. Intern. Telemetering Conf., Vol. 14, 14-16 November 1978, Los Angeles, CA, pp. 611-618.

## SECTION 3

### SYSTEM IMPLEMENTATION ANALYSIS

A variety of methods and techniques can be used to implement the various elements of a satellite communications system. The implementation approach used for these elements can have a significant impact on the structure and architecture of the on-board processor. This section reports the results of studies to identify a possible system control procedure, to analyze multiple access techniques and modulation methods, and to analyze baseband regeneration performance. In addition, an analysis is presented showing the relationship among various parameters of the system and the processor.

#### 3.1 SYSTEM CONTROL

The wideband satellite system described in this report must be capable of operation over a wide range of user traffic characteristics, e.g., user response time, speaker-to-listener turnaround time, blocking probability, etc., and be compatible with Bell System's future Common Channel Interoffice Signaling (CCIS). Satellite resource allocation can be achieved by dedicated, demand, or hybrid assignment techniques. Demand Assignment Multiple Access (DAMA) techniques depend upon the system control adopted for interconnectivity and efficient channel utilization. The system control approach can be either central or subcentralized.

In this section, after traffic characteristics and relevant aspects of CCIS are briefly described, the major emphasis is on a satellite overlay configuration and communications protocols. The configurations of the subsystems, e.g., users, regional control center (RCC), processing satellite with on-board controller, and ground based system control and management center (SCMC) are defined and their functions are explained. Various access schemes employing a control channel (orderwire) are identified.

Two satellite system overlay protocols, (a) a call establishment procedure which is circuit switched and central controlled, involving a double hop protocol; and (b) a destination variable procedure which is packet switched and subcentralized, involving a single hop protocol are discussed. The performance of the call establishment protocol in terms of number of circuits and utilization for given blocking probabilities is analyzed. The

delay-utilization performance analysis for different numbers of channels and path lengths is initiated for the destination variable protocol. Examples include fixed beam and scanning beam cases.

Although system control issues do not govern the technology, analysis and modeling are necessary for system design guidance and evaluation. Recommendations and aspects of system control and management which require further analysis and simulation are included.

### 3.1.1 Traffic

The present study primarily addresses future satellite communications which will off-load ground trunking networks ( $T1 = 1.544$  Mb/s,  $T2 = 6.312$  Mb/s,  $T3 = 44.736$  Mb/s, and  $T4 = 274.176$  Mb/s) carrying a wide variety of services with different transmission characteristics. The traffic is classified into three major categories: real time (voice, video), near real time (interactive, query/response) and non real time (bulk data traffic, remote job entry).

The satellite system must be compatible with not only the existing terrestrial links (wire lines, line-of-sight radio, and microwave links) but also the future intercontinental service. A satellite processing network node can assist in adaptability and interoperability in the meantime.

Real time traffic is characterized by relatively long continuous transmissions which are either accepted, blocked, or scheduled by system control. The circuit capacity is allocated for the duration of the session.

Near real time traffic is often bursty and characterized by frequent, short messages such as interactive data requiring immediate delivery; whereas long messages of the store-and-forward type do not require immediate delivery.

Since non real time data (e.g., low priority jobs submitted for overnight processing) require no immediate delivery, more delay can be tolerated than with near real time traffic.

### 3.1.2 Terrestrial Signaling Schemes

Wideband satellite services compatible with the Bell System's future CCIS are desirable. Relevant aspects of the CCIS are

described in this section [Bohacek, 1978; Nowack, 1978; Frerking, 1978]. Interoffice signaling initiates and terminates connections, and communicates call status and addresses. Historically, such control information has been carried on voice trunks. The recent CCIS uses a physically separate channel for signaling purposes.

The early automatic switching systems employed d.c. signaling techniques where the supervisory information is communicated by level or direction of current flow. The later carrier systems precluded d.c. signaling and used a.c. signaling. Single frequency (SF) a.c. signaling systems require signaling equipment to be permanently connected to both ends of each voice trunk. Multifrequency (MF) a.c. signaling communicates addresses between switching offices using combinations of two out of six tones in the voice-frequency range to represent ten digits and up to six control signals. SF and MF systems are widely used for both analog and digital transmissions.

Since May 1976, the CCIS toll network has been used for signaling associated with direct distance dialing (DDD) calls. By 1985 about 80% of intertoll trunks will be provided with CCIS signaling capabilities. This will reduce call setup time and improve signaling reliability. Also, during 1980, it is planned to extend CCIS to new services like Traffic Service Position Systems (TSPS), automating credit-card calls, collect calls, and bill to third party calls by accessing a data-base to verify customer dialed information [Dahlbom, 1978].

Some of the significant features of the CCIS, a packet switching network, are:

- Data rates
  - 2400 b/s (at present)
  - 4000 b/s, 64 kb/s (future)
- Faster signaling than a.c. or d.c. schemes, e.g., call setup times are 1-2 sec as opposed to 10 sec for today's service.
- More complete connectivity.
- Higher flexibility, e.g., signaling can occur at any time in either direction even when a call is in progress.
- Immune to co-channel interference because separate signaling channels are employed.



- Less expensive because dedicated signaling equipment is no longer required.

Figure 3-1 depicts a CCIS network consisting of an arbitrary number of switching offices (SOs), several types of links, and two signal transfer points (STPs) per DDD region, which perform message routing functions. A satellite overlay is also shown in this figure. Each STP is connected to several switching offices by A links. A links are always provided in pairs, one link to each STP. The STPs in a given region are interconnected to the STPs in all other regions by B links. These links are always provided in groups of four between any two regions. The two STPs in a region are connected by C links that carry update and status information and serve as alternate paths for other signaling traffic during failure modes. These can be an arbitrary number of SOs. Each SO is connected to both regional STPs by a CCIS link. An STP routes an incoming CCIS message to the appropriate A link if that message is for a switching office in the same region or routes it to an STP in the destination region over a predetermined link according to the local routing tables.

Under current operation, a given telephone trunk is associated with a specific signaling path. A call between two switching offices to be sent over a connecting trunk is set up via a fixed path in the separate CCIS network using predetermined local assignment tables at the switching offices and the STPs [Frerking, 1978; Dahlbom, 1978]. The switching office at the source selects an outgoing trunk and transmits the appropriate signaling message to a predetermined STP. The STP uses the local assignment tables which uniquely associate incoming labels with outgoing labels. Since both end points of every telephone trunk are predetermined, the routing of signals associated with the trunk can be fixed.

For new customer services that require direct links (shortest path in the CCIS network) between switching offices which do not share any direct trunks, a global routing table is required. Maintaining the current routing information at each switching office would be an enormous task because there are too many switching offices. For example, No. 1 ESSs alone will number about 2000 by 1980 [Dahlbom, 1978]. Instead, if the routing tables are provided at the 20 STPs (10 regions and two STPs per region), two orders of magnitude of memory saving can be obtained.

The CCIS signaling scheme provides the capability of routing a message based entirely upon address information. A message could be routed to the destination switching office or a STP node with a data base (which has a list of a large number of possible addresses but

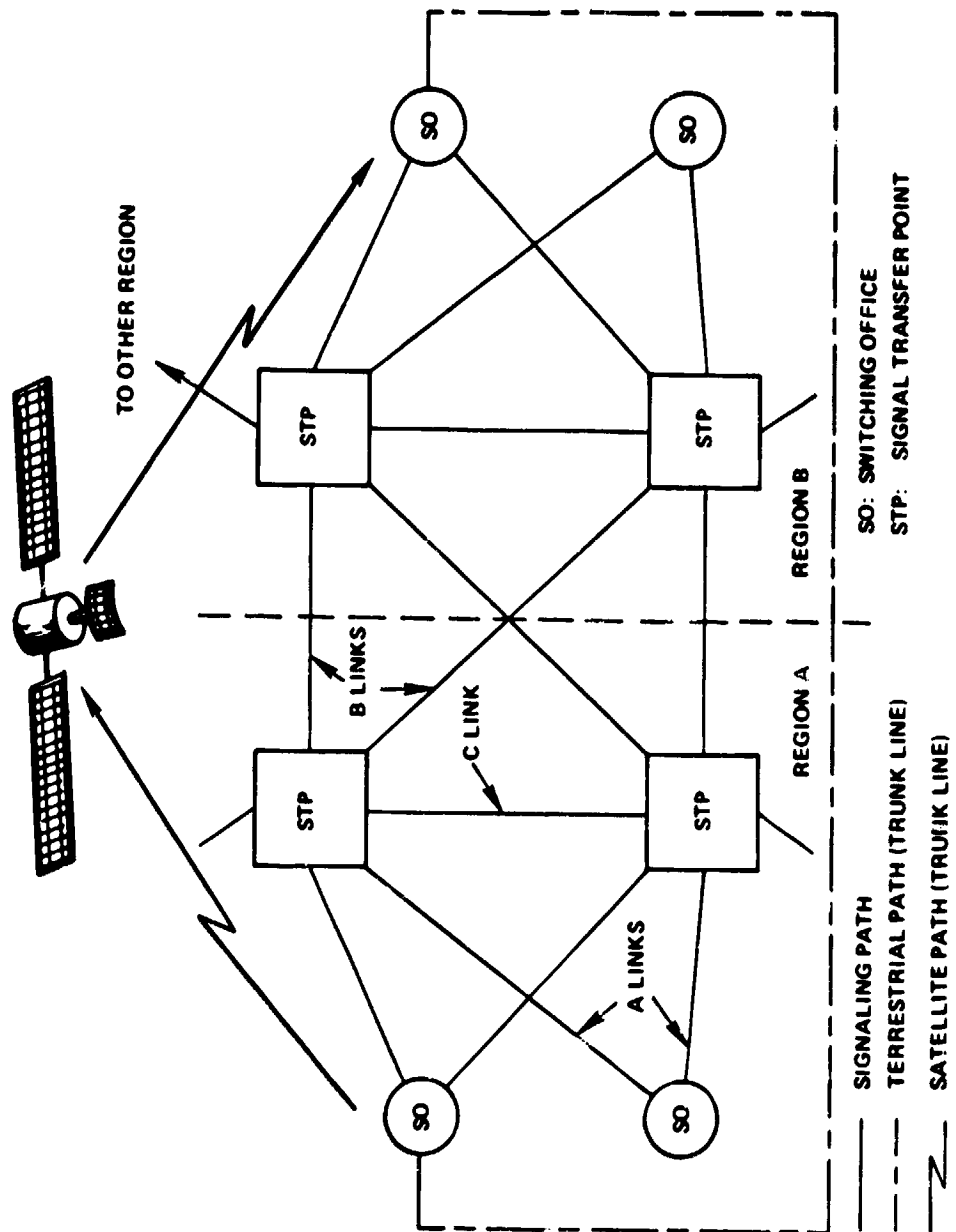


Figure 3-1. Common Channel Interoffice Signaling (CCIS) With Satellite Overlay

not a switching office node) containing customer service information. An STP can have a data base of information which can be updated remotely by a switching office. Such a data base can have a wide ranging utility in providing convenient service to users. For example, if a person is going to be away from the office, incoming phone calls could be automatically routed to a temporary new number prestored in the data base.

### 3.1.3 Variable Assignment Techniques

Assignment procedures for allocating satellite resources to system users can be either dedicated or on demand. A combination of fixed and demand assignment is also possible. As a result, satellite channel utilization can be improved, satellite power can be used more efficiently, and the system can be dynamically adapted to traffic fluctuations and various priorities.

DAMA can be classified into two generic schemes, fully variable and semivariable, in accordance with the way source and destination channel connections are made. In a fully variable demand assignment (DA) all channels are pooled [Hilborn, 1977] so that any user may use any channel according to the instantaneous traffic load. For a given blocking probability, the number of channels required to transfer a given amount of traffic in a fixed assignment system is greater than in a fully variable system.

If users are limited to sending information over a unique set of channels and receiving on any channel the system is known as destination variable demand assignment. A system in which users receive only on specific channels and transmit on any channel is called source variable demand assignment.

In a central control demand assignment scheme, user requests are transmitted by regional centers over satellite subchannel(s) of limited capacity and the central controller arbitrates requests for capacity allocation, returns control channel information, and establishes the communication links. The centralized controller is assumed to have comprehensive information about the total network state. Accordingly, the network organization and allocation of capacity is expected to be quite efficient.

A subcentralized control philosophy can provide a faster user response time than purely centralized control. In addition, it may be advantageous to distribute more of the real time control processing to the satellite and regional centers. Both centralized and subcentralized control schemes are considered in this report.

#### 3.1.4 Satellite Overlay Configuration

Figure 3-2 depicts an overall network configuration consisting of different types of user (voice, video, and data), regional control centers (RCCs), a processing satellite with an on-board controller, and a ground based system control and management center (SCMC). The SCMC consists of a system control center (SCC) and a system management center (SMC). For demand assignment purposes, users within a region form a star network centered around an RCC. Requests are forwarded by an RCC to the SCC using a control channel and the assignments or any response messages are transmitted back to the concerned RCC. The various constituents of the network and their functions are discussed in this subsection.

Regional Control Center. The continental United States (CONUS) can be partitioned into regions which each have a regional control center. Each RCC provides access to the regional users by forwarding their requests for service and translating commands from the SCMC.

An RCC principally serves as a coordinating and processing node for regional users. For the satellite overlay, it is hypothesized that every RCC interfaces with the terrestrial CCIS network as well as the satellite for control signaling and message transmission purposes.

On-Board Controller. The detailed implementation of the on-board control configuration depends on the accessing and downlink multiplexing scheme assumed. The on-board controller (OBC) can be thought of as a microprogrammed processor. It supervises the reorganization of the uplinks for the downlinks. It processes the requests for service with a predetermined control algorithm and interprets the ground based SCC commands. The selection algorithm can be either purely random (a choice of controller) or based upon priority of the user transmission. The controller can receive data base information about the requests, priorities, and assignments to assist the selection algorithm. The demodulated data from uplinks is demultiplexed and switching is accomplished at baseband, i.e., the addresses which associate uplinks are routed to the proper downlinks by the controller. The digital processor is treated in more detail in section 4.

Some of the functions of the on-board controller include:

- Providing a satellite interface between the regional control centers and SCMC

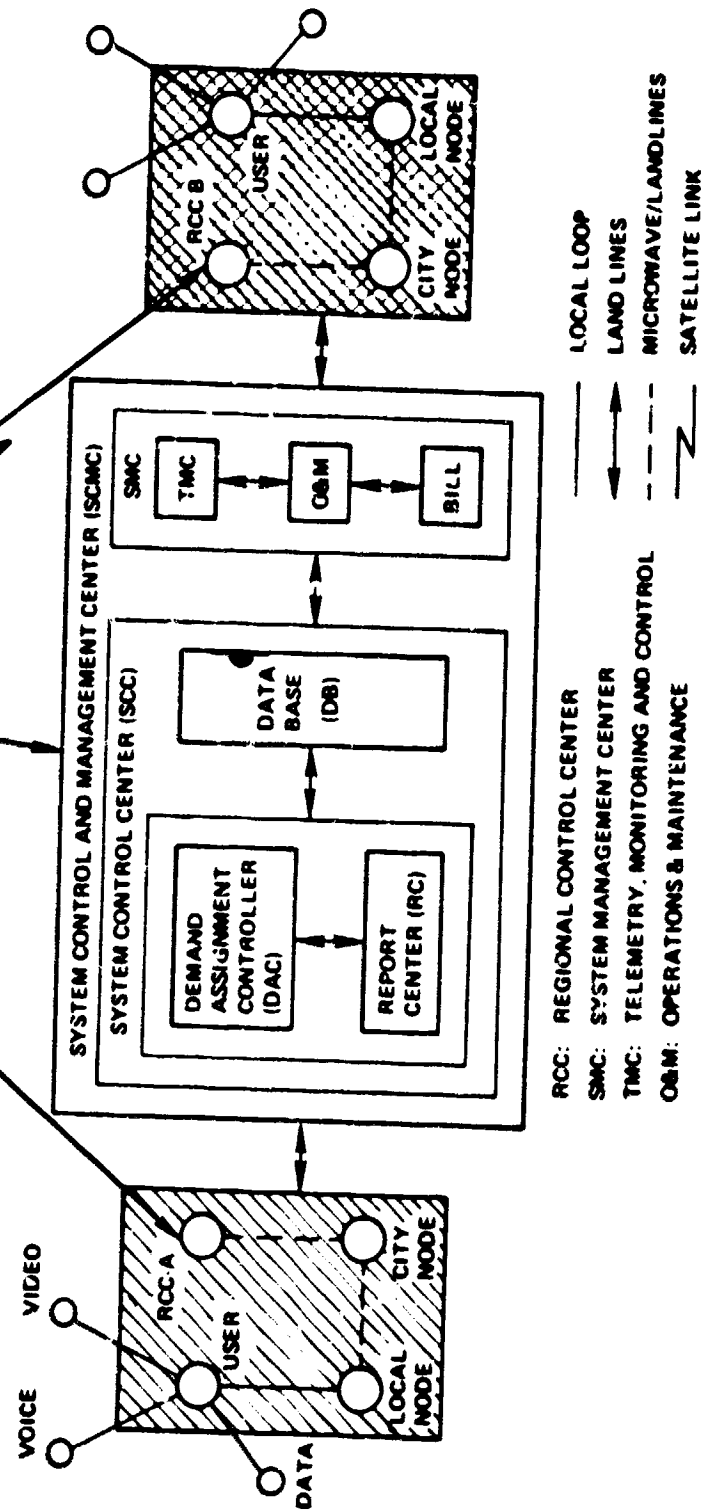


Figure 3-2. Overall Network Configuration

- Arranging transmultiplexing among different user services
- Uplink demultiplexing, buffering, address recognition, formation of superheaders, packetizing to standard formats; downlink selection and transmission/broadcasting to addresses if necessary
- Implementing flow control mechanism, error control and recovery procedures as required to control the information flow from the regional control centers and pass status information to the RCCs and SCMC

System Control and Management Center. As shown in figure 3-2, the SCMC consists of two principal constituents: an SCC, and an SMC. The SCC and the SMC are discussed in this subsection.

The SCC allows point-to-point, broadcast and conferencing modes of communication. The automated procedures permit relatively simple demand assignment protocols for the assignment of circuits/slots to users and release of the circuits/slots on completion of the message transmission. If priority processing is also assumed in the configuration, a high priority user transmission (even on a packet basis) can be successful by preempting the low-priority ones.

The SCC consists of a demand assignment controller (DAC), in which adaptive traffic algorithm and assignment algorithm(s) reside, a report center, and a data base (DB) as shown in figure 3-2. The SCC interfaces with the SMC, the RCC, and the OBC. The SCC essentially controls the system resource sharing and assignments according to the requests of RCCs. It also performs system monitoring continually and provides status information to active users via an RCC.

The DB contains all the necessary information for the assignment algorithm. The various lists of the DB include: users/RCCs/addresses, requests/assignments/preemptions, type/priority/delay requirements, current usage of resources, and capacity available.

The addresses include the incoming and outgoing users as well as the RCCs; the type of user and priority information is maintained along with the delay tables as a separate list. The current usage information include the slots, channels or circuits assigned and time of assignments. In addition to these, the request list includes an input queue indicating the order and priority of arrival. Whenever an assignment is made to a user the request is

dropped from the list. The table also indicates the available capacity to every regional center enabling it to make requests automatically to the SCC.

An adaptive traffic algorithm updates the data base based upon the information about the system parameters, link parameters, network congestion, and the telemetry monitoring data. Secondly, it preempts the low priority user under system overloading conditions and informs the concerned user.

The DAC obtains all the relevant information on the parameters, requests and available system capacity from the DB. Using the data base information, an assignment algorithm makes circuit/channel/slot assignment to the RCCs. Also, it prepares control bursts regarding the response messages, e.g., adaptive system control parameters, user preemptions, and any broadcast information to the users. Some of the major functions of the DAC are: RCCs data buffering, request processing/assignments/preemptions, error checking and retransmissions schemes, and continuous command and telemetry data exchange between the satellite and the controller.

The report center obtains requests from the RCCs, link quality parameters, and traffic status on a near term basis. It supplies the various lists/reports and updates to the data base and DAC enabling the assignment algorithm to successfully sanction the requests.

The SMC will act as a coordinating, monitoring and reporting element of the system. For unattended system operation, the SMC will provide monitor, fault detection and control.

The SMC consists of a telemetry monitoring center (TMC), operations and maintenance (O&M) center and billing center (BC) as shown in figure 3-2.

The TMC subsystem collects and processes telemetry data to monitor satellite performance and operating modes. When a failure or abnormal condition occurs, the TMC performs diagnostic tests and takes corrective action.

The O&M center is a data processing facility which provides maintenance support to the RCCs via automated system diagnostics. A variety of administrative support functions are also furnished.

The BC maintains a record of the system usage and provides user billing/accounting information to all RCCs. The billing can be

based in part upon usage statistics obtained by the on-board controller.

### 3.1.5 Access Schemes

The channels are of two types: message and control. RCC access to a message channel is handled via a control channel. The message channel carries the user data from source to destination once the assignment has been made by the SCC.

The control channel carries RCC requests for capacity, i.e., communication circuits, slots in a TDMA organization or channels in a FDMA organization, and subsequent assignments from the DAC of the SCMC. It can also be used to broadcast system status messages using data base or SMC information, to the regional center.

Some of the characteristics of a control channel can be assumed to be: a physically separate channel from the message channel, relatively low data rate (e.g., 2400 b/s). The segment of the control channel between SCMC and the satellite should operate at S-band to avoid outages due to rain.

Two fixed and random accessing techniques for the control channel are identified. A reservation assignment with a fixed assignment control channel provides exclusive capacity to each user to make requests for message transmission capacity. A specific amount of capacity will be reserved for that user by the SCC.

The reservation-random assignment scheme is different from the previous one, in that there is contention for the reservation making capacity on the control channel. In a random access scheme, the control channel is shared by a large number of users, in a completely unsynchronized manner, on a contention basis with other users [Abramson, 1977]. If two user packets arrive at the SCMC at the same time a collision or conflict occurs. An operational protocol can reduce the number of contentions.

Based on some of the existing analyses [Kleinrock, 1977] it can be seen that reservation-random and reservation-fixed techniques are superior to polling schemes with regard to the delay performance. Hence, these two reservation techniques are recommended for the system under study and also for future extensive analysis and simulation studies.



### 3.1.6 Communication Protocols

After a brief description of the CCITT X.25 interface for data transfer and a few line control procedures are listed, the bulk of this subsection is devoted to the two main satellite overlay protocols addressed in this report.

In order to provide an efficient means by which a large number of diversified terminals can gain access to the packet switching networks, the International Telephone and Telegraph Consultative Committee (CCITT) has recommended that the international packet mode interface be X.25. Satellite system planners and designers should be cognizant of any essential standards or procedures for defining compatible satellite overlay protocols.

The X.25 interface consists of three distinct levels of independent control procedures listed below:

- Physical level interface
- Frame level interface
- Packet level interface

The physical level interface specifies the use of a duplex, point-to-point synchronous circuit, thus providing a physical transmission path between the terminal and the network.

The frame level logical interface specifies the use of a data link control procedure which is compatible with the high level data link control (HDLC) procedures being standardized by International Organization for Standards (ISO).<sup>\*</sup> The use of this data link

---

<sup>\*</sup>Other link control procedures (line protocols) are of two types: (a) character oriented, e.g., binary synchronous communications (BSC) or BISYNC, originated by IBM in 1968, and Digital Data Communications Message Protocol (DDCMP) originated by DEC; (b) bit oriented, e.g., advanced data communication control procedures (ADCCP) originated by American National Standards Institute (ANSI) in 1973-74 which is recommended for computer to computer or computer to terminal over common carrier, satellite link or dedicated line; synchronous data link control (SDLC) originated by IBM in 1973 which is the same as ADCCP but with minor differences (also recommended for synchronous transmission for computer to computer or computer to terminal over common carrier, satellite link, or dedicated line), and control data communications control procedure (CDCP) originated by Control Data Corporation (CDC) in 1975.

control procedure ensures that packets provided by the packet level and contained in HDLC information frames are accurately exchanged between the terminal and the network. The frame level interface provides the packet level logical interface with an error free variable delay link between the terminal and the network.

The packet level logical interface, the highest level of X.25, specifies the set of conventions governing the manner in which packet terminals format control information and data into packets, establish, maintain, and clear calls, and manage the flow of data to and from the network.

#### 3.1.6.1 Call Establishment Protocol

A call establishment protocol which is based upon circuit switching techniques is described in the following paragraphs. The protocol uses a central controller for user demand assignment involving a double hop via the satellite as depicted by the packet sequence in figure 3-3. The termination of a call can be handled similarly.

Users are assumed to generate an average of  $\lambda$  requests per second according to a Poisson process. The duration of a call is assumed to have an exponential distribution with an average call duration of  $1/\mu$  seconds, so the offered traffic load is  $a = \lambda/\mu$  (Eriangs). A total of  $s$  circuits is assumed.

If a circuit is available the SCC will assign it to the requesting user. However, if a request arrives when all the circuits are busy, it can be either treated as lost (blocked call cleared model) or buffered until a circuit is available (blocked call delayed model).

Blocked Calls Cleared. If  $k$  is the number of busy circuits, a birth and death model [Kleinrock, 1976] is applicable to the present situation:

$$\lambda_k = \begin{cases} \lambda, & k < s \\ 0, & k \geq s \end{cases} \quad (3.1)$$

$$\mu_k = k\mu, \quad k = 1, 2, \dots, s. \quad (3.2)$$

( $\mu_k = k\mu$  because if there are  $k$  calls in progress the probability of any call terminating is  $k$  times that of one call terminating.)

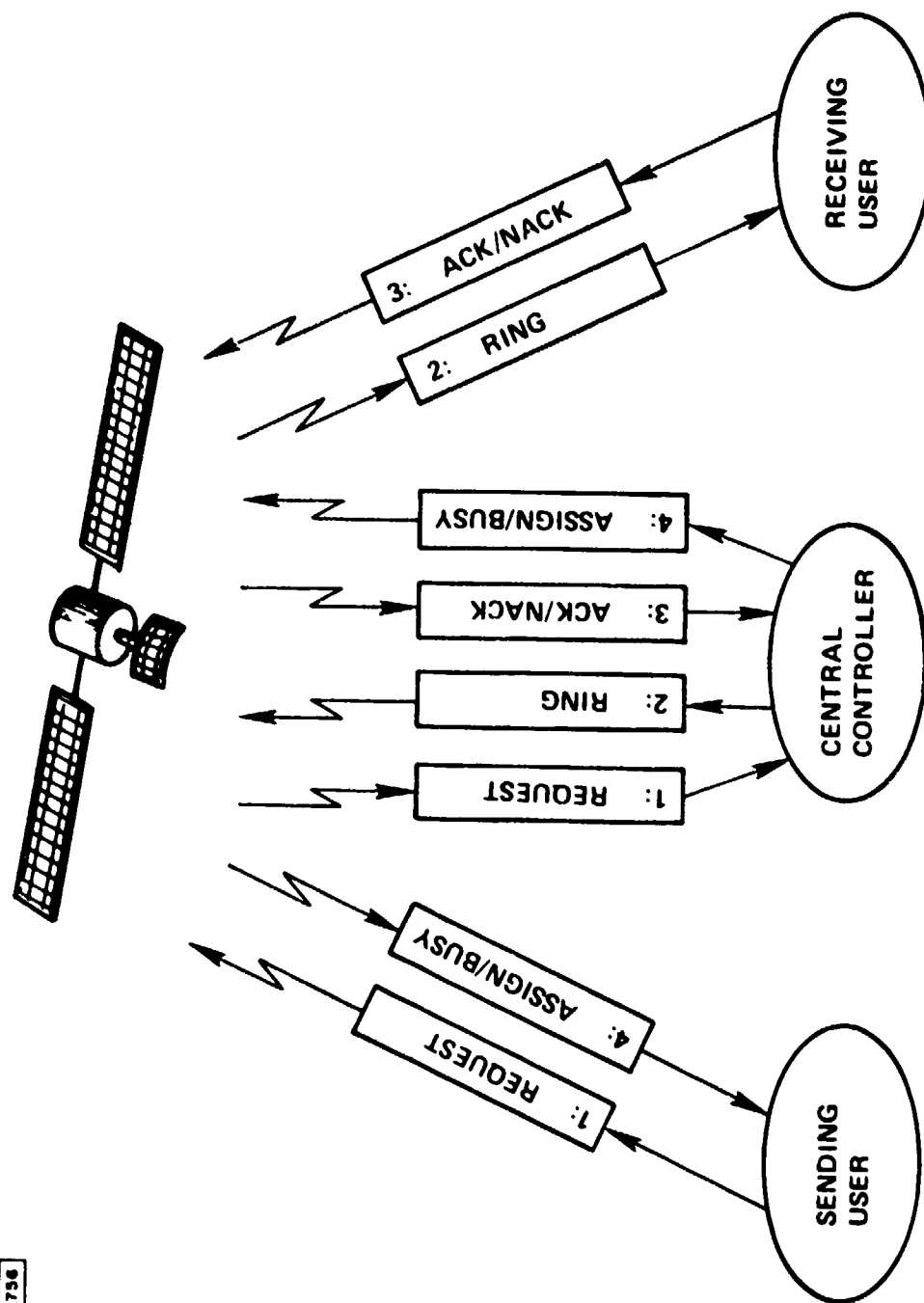


Figure 3-3. Call Establishment Protocol

The probability that the system contains  $k$  users at an arbitrary time  $t$  is given by

$$p_k = \begin{cases} p_0 a^k / k!, & k \leq s \\ 0, & k > s \end{cases} \quad (3.3)$$

where

$$p_0 = 1 / \left( \sum_{k=0}^s a^k / k! \right). \quad (3.4)$$

The probability that all  $s$  circuits are busy, i.e., the blocking probability, is given by

$$p_s = p_0 a^s / s! = B(s, a) \quad (3.5)$$

$B(s, a)$  is called Erlang's B formula or loss formula.

In the blocked calls cleared mode, the offered load is greater than the carried load. The carried load  $a'$  is defined as the average number of busy circuits [Cooper, 1972]:

$$a' = a[1 - B(s, a)] \quad (3.6)$$

the product of the offered traffic and the fraction of the requests that are not blocked.

The system utilization efficiency, or load carried per circuit, is given by

$$\rho = a'/s < 1. \quad (3.7)$$

Figure 3-4 shows curves of the number of circuits vs. utilization for selected values of blocking probability. For the present application, the number of circuits equals the number of beams times the number of carriers/beam.

#### Example 1: Fixed Beams

If there are 20 fixed beams and two carriers/beam, the total number of fixed beam circuits is  $s=40$ . Assume the blocking probability  $B(s,a)$  to be 1% according to CCITT recommendations [CCITT, 1965]; figure 3-4 provides the system utilization of 71.5%. At 90%, a blocking probability of 0.163 is implied.

#### Example 2: Scanning Beams

If there are four scanning beams with two carriers/beam and two scanning beams with one carrier/beam, the total is 10 circuits ( $s$ ). Assume a blocking probability of 1%; this provides a system utilization of 44%. A blocking probability in excess of 0.4 is implied at a 90% utilization.

Blocked-Calls Delayed. Consider the case in which the user requests are buffered once all  $s$  circuits are busy, i.e., blocked-calls delayed. The system is considered with infinite buffering (for theoretical analysis). The request arrival process is Poisson and the message lengths are exponential. This model is known as an  $M/M/s$ \* queueing system. The circuits are assigned in the order they arrive. Again, the birth and death formulations provide:

---

\* $M/M/s$  - exponential interarrival distribution/exponential service time distribution/number of circuits.

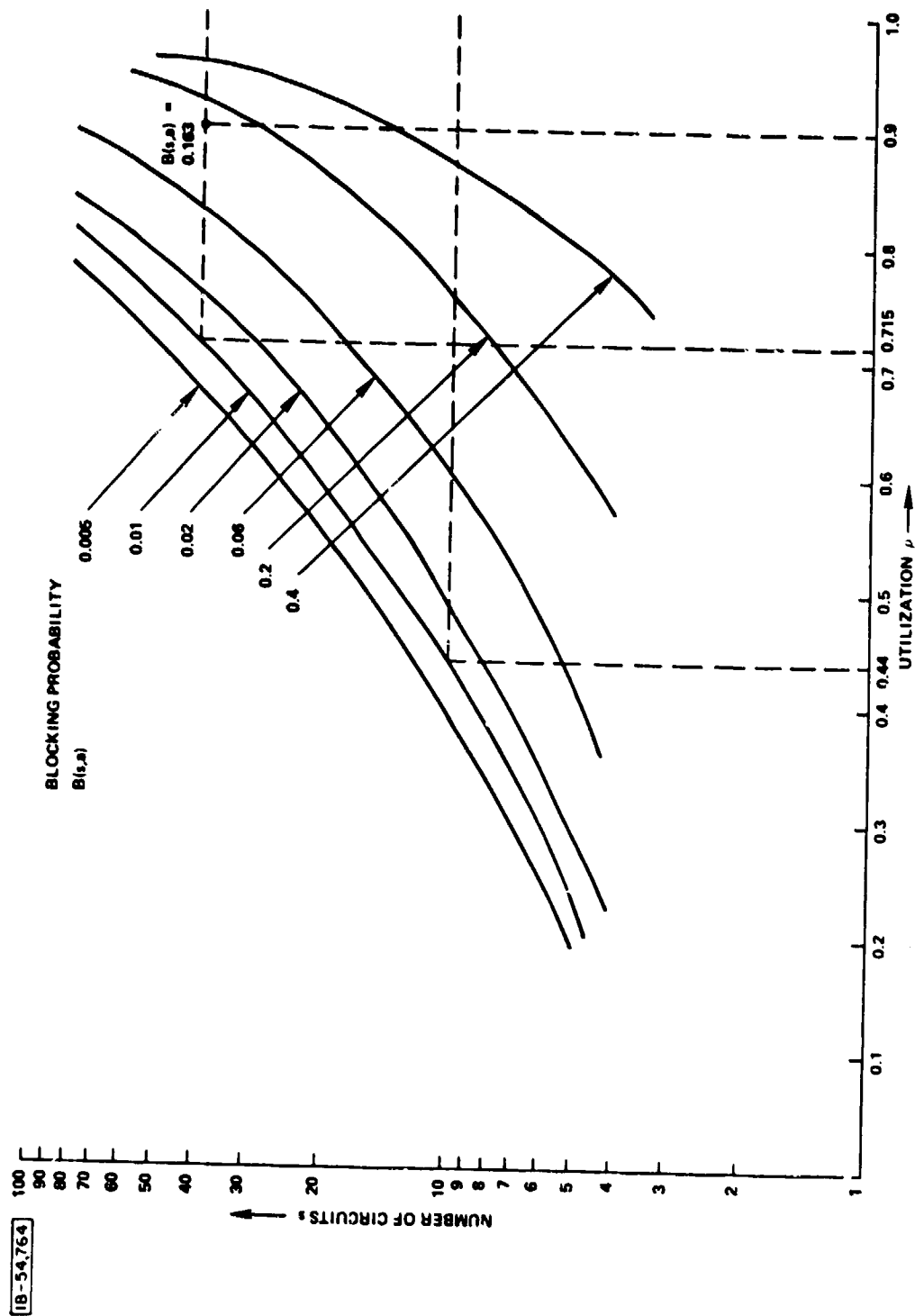


Figure 3-4. Blocked Calls Cleared Model Performance

$$\lambda_k = \lambda, \quad k = 0, 1, 2, \dots \quad (3.8)$$

$$\mu_k = \begin{cases} k\mu, & 0 \leq k \leq s \\ s\mu, & k \geq s \end{cases} \quad (3.9)$$

The solution of  $p_k$  can be given by [Kleinrock, 1976]

$$p_k = \begin{cases} p_0 (\rho s)^k / k!, & k \leq s \\ p_0 \rho^k s^s / s!, & k \geq s \end{cases} \quad (3.10)$$

where  $\rho = \lambda / (s\mu) = a/s < 1$  and

$$p_0 = 1 / ((1/(1 - \rho)) (a^s / s! + \sum_{k=0}^{s-1} a^k / k!)). \quad (3.11)$$

The probability that the arriving request is forced to join the queue can be derived to be

$$P(\text{queuing}) = \sum_{k=s}^{\infty} p_k = a^s / (s! (1 - \rho) p_0) = C(s, a). \quad (3.12)$$

Equation (3.12) is called Erlang's delay formula denoted by  $C(s, a)$ . In this case the carried load will be equal to the offered load, i.e.  $a' = a$ . In the Erlang delay model all blocked requests wait as long as necessary for assignment, and therefore all user requests are serviced.

Figure 3-5 describes Erlang's delay formula  $C(s, a)$  plotted against offered traffic  $a$  in Erlangs for different values of the number of circuits  $s$ .  $C(s, a)$  provides the portion of requests that find all the circuits are busy and consequently are buffered until the assignments are made.

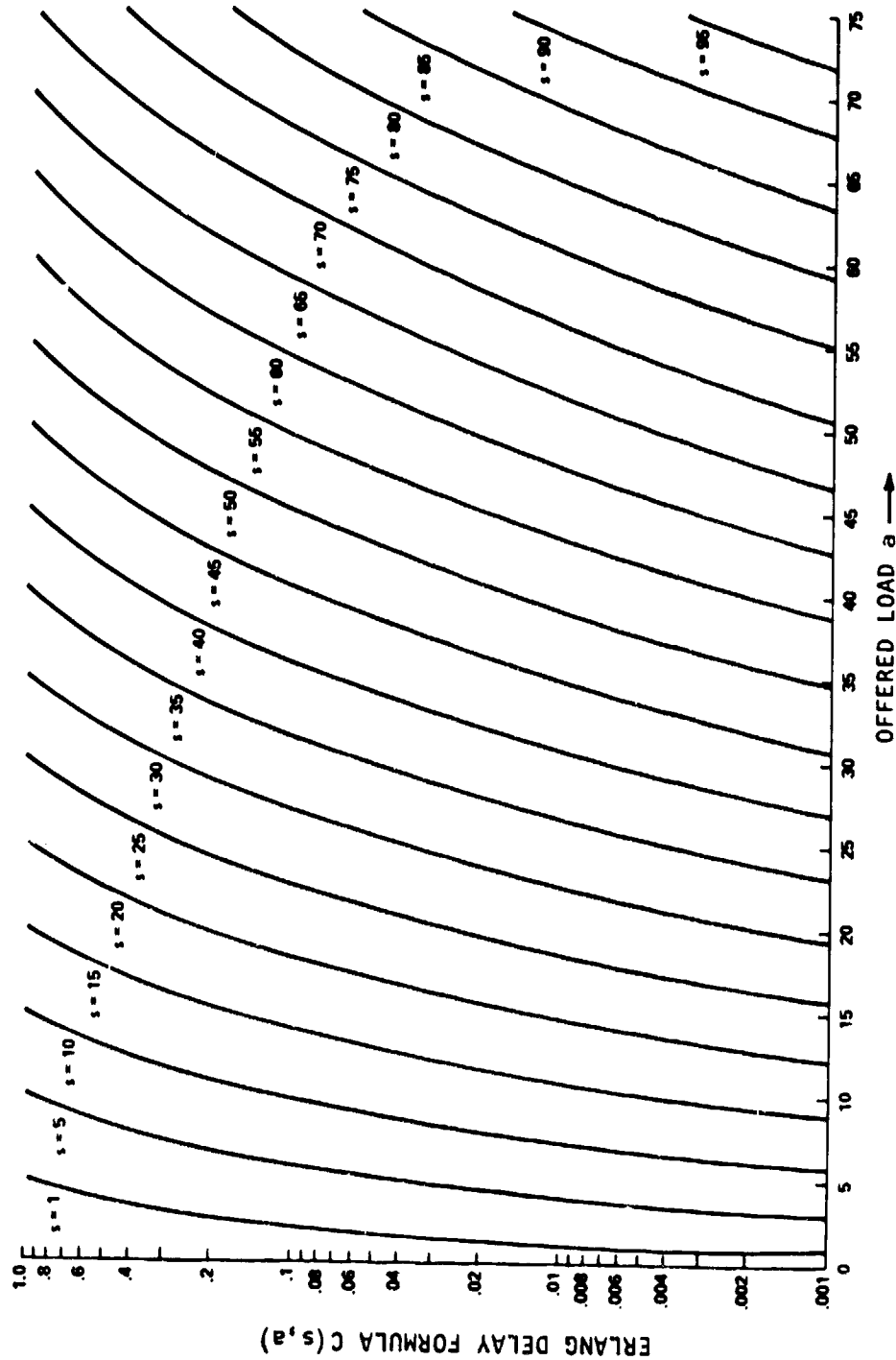


Figure 3-5. Erlang Delay Formula  $C(s, a)$  vs. Offered Load  $a$  (Erlangs)



The waiting time distribution function for the Erlang delay model (Poisson input and exponential service time) is as follows:

Let  $W$  be the waiting time or delay experienced by an arbitrary request for assignment. The conditional waiting time distribution is given by [Cooper, 1972]

$$P(W > t | W > 0) = \exp(-(1 - \rho)s\mu t) \quad (t \geq 0, \rho < 1) \quad (3.13)$$

Figure 3-6 presents a plot of  $P(W > t | W > 0)$  against  $s\mu t$  for selected values of  $\rho$ . The unconditional waiting time probability  $P(W > t)$  is given by

$$P(W > t) = C(s, a) P(W > t | W > 0). \quad (3.14)$$

The conditional mean waiting time  $E(W | W > 0)$ , the waiting time experienced by the requests which are blocked is given by

$$E(W | W > 0) = 1 / ((1 - \rho)s\mu) \quad (3.15)$$

The overall mean waiting time  $E(W)$  is given by

$$E(W) = C(s, a) E(W | W > 0). \quad (3.16)$$

In order to calculate the probability of waiting more than a specified time for message transmission and the overall average mean waiting time for the blocked-calls delayed model, the following parametric values are assumed.

Example 1 (Fixed Beams)

Let

number of circuits,  $s = 40$

utilization (design goal),  $\rho = 80\%$

waiting time range,  $t = 1$  to 10 sec

message length,  $1/\mu = 10-100$  sec

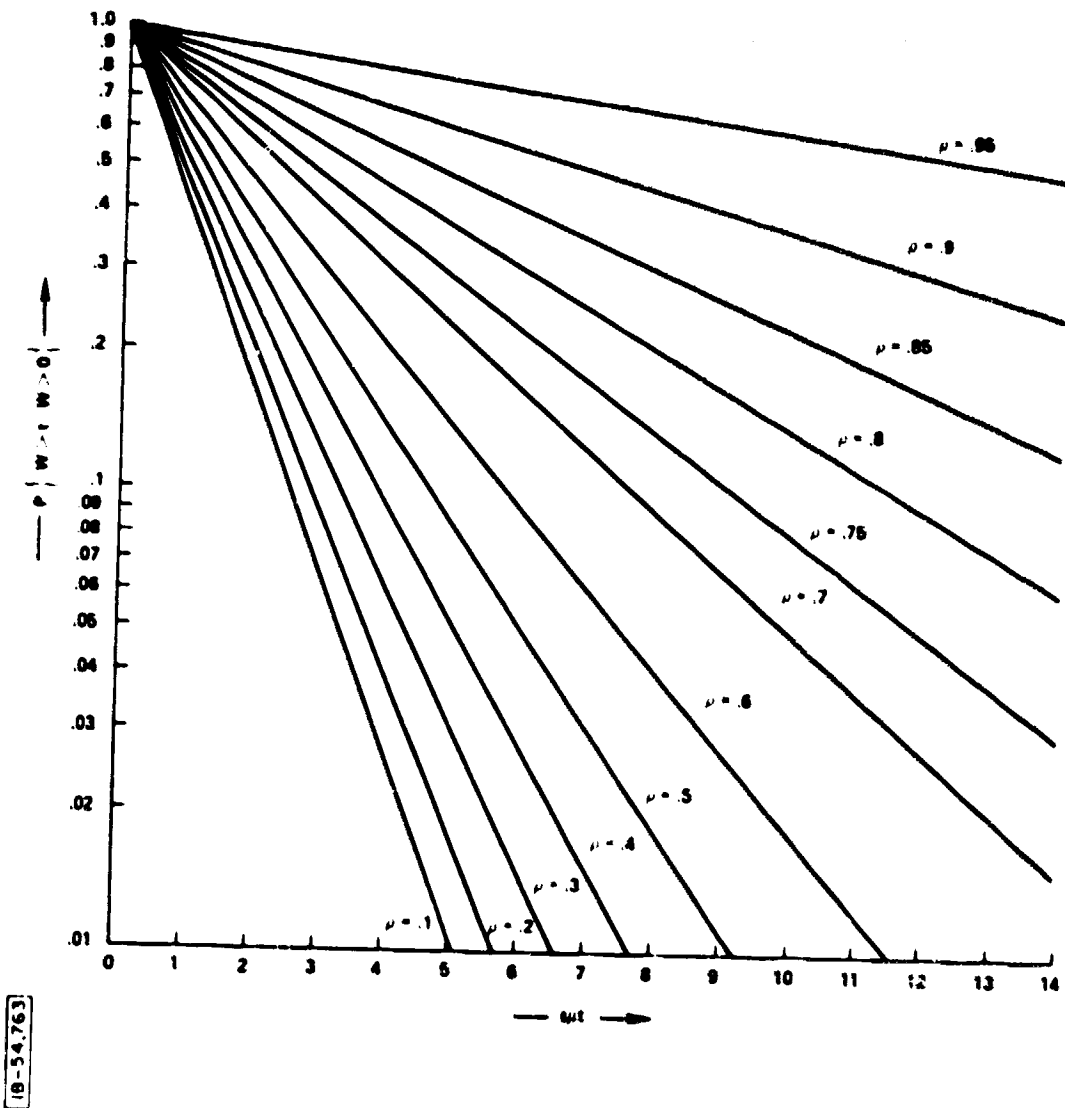


Figure 3-6, Conditional Probability for a Blocked - Calls Delayed Model

Since the blocked requests are delayed until the assignment is made,  $C(s,a)$  is known as the Erlang's delay formula. However,  $C(s,a)$  provides the probability that the circuits are busy.

From figure 3-5 the blocking probability  $C(s,a)$  for the given  $s$  and  $\rho$ , ( $\rho = a/s$ ) is found to be 0.02. The conditional probability  $P(W>10|W>0)$  for an average message length of 100 sec is obtained from figure 3-6 to be 0.46. The unconditional probability  $P(W>t)$  that a user message waits beyond 10 sec for transmission is therefore 0.0092 from equation (3.14). This increases to 0.0132 for  $t=5$  sec.  $P(W>t) = 0.004$  for  $t=10$  sec and  $1/\mu = 50$  sec.

The overall mean waiting time  $E(W)$  is 0.25 sec from equation (3.16).

#### Example 2 (Scanning Beams)

Let

number of circuits  $s = 10$

utilization  $\rho = 80\%$

waiting time  $t = 10$  sec

message length  $1/\mu = 100$  sec

The probability of waiting more than 10 sec is 0.017 and the overall mean waiting time is 1 sec.

#### 3.1.6.2 Destination Variable Protocol

The destination variable demand assignment protocol discussed below is based upon packet switching concepts and employs a single hop over a satellite. In addition, a regenerating satellite performs a certain degree of assignment and scheduling with the help of the uplink packet header information (e.g., packet length, type, priority, etc.). Considerable delay can be eliminated by routing retransmitted (unsuccessful) message packets due to link/data errors over the terrestrial network.

Figure 3-7 describes simultaneous point-to-point and broadcast modes of operation via TDMA on the uplink and FDM/TDM on the downlink.

The hybrid network model depicted in figure 3-2 consists of the following: a set of nodes interconnected by a terrestrial channel, including a set of regional control centers colocated at earth stations, and a regenerating satellite linking all RCCs and the central controller (SCMC).

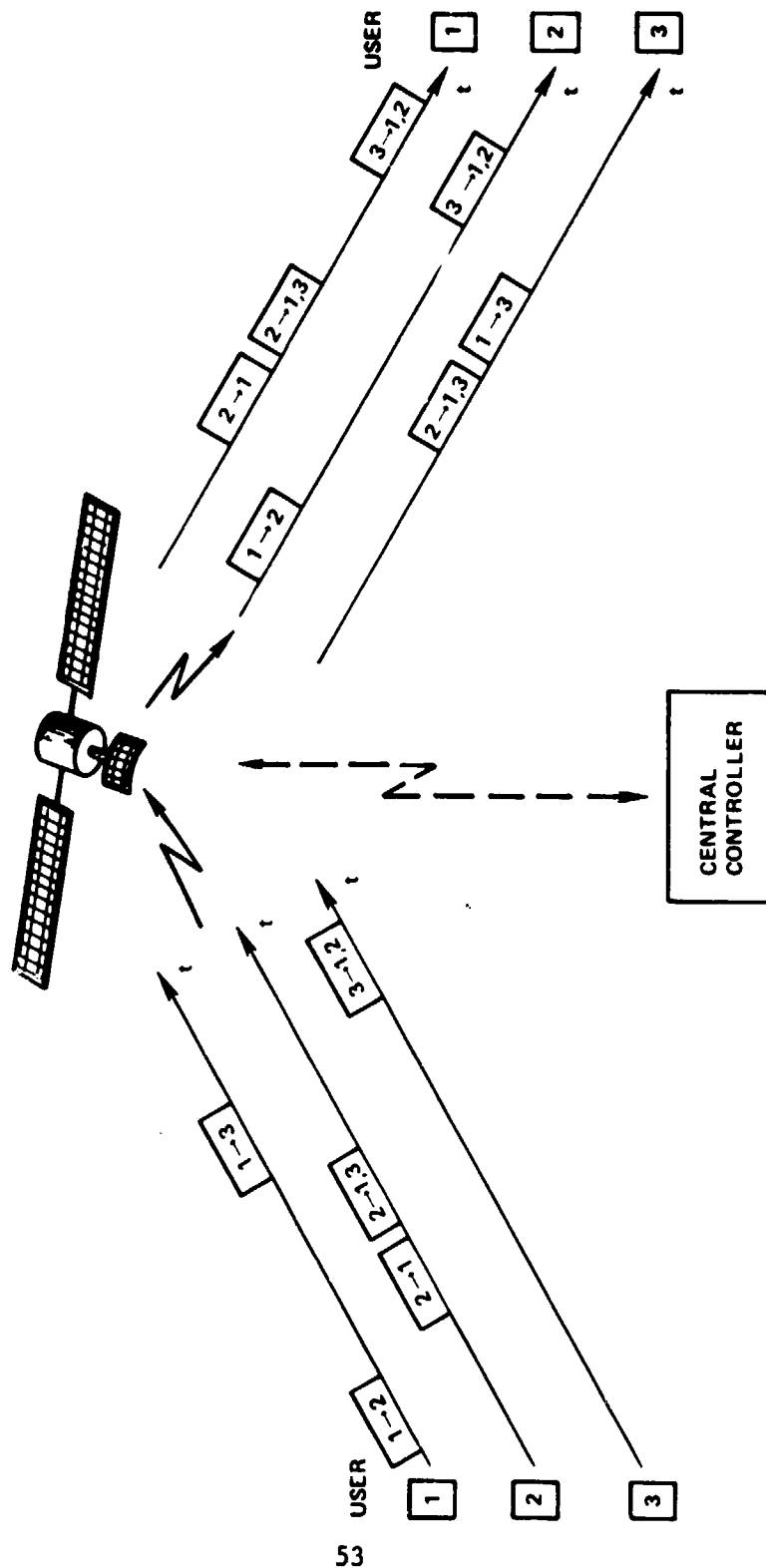


Figure 3-7. Destination Variable Protocol

The following paragraphs describe the protocol operation and an end-to-end delay analysis for the proposed network. A few examples are included.

The protocol operation is as follows:

- A user transmit the message packets with address information to the RCC via local office connections. The RCC collects data packets from different users within the region and forwards them to the on-board controller along with the regional header and identification number. The RCC also examines the data for errors and provides necessary protection on a regional level.
- The on-board controller processes the uplink header, examines the message type, priority, acceptable delay information, and error check.
- The controller prepares a downlink super header to different RCCs and assigns packets to downlink channels. (No central controller involvement is assumed for assignments.)
- If the packets of some users are not successful (e.g., the users do not receive the assignments/acknowledgment within certain time out period) the retransmissions are rerouted via terrestrial network.
- The on-board controller broadcasts any commands or status messages regarding the current assignments/pre-emptions, etc., once every frame and keeps updating the same.
- An RCC selects downlink packets destined for its region and distributes individual packets according to local user addresses contained therein.

In store-and-forward packet switching, better utilization for data traffic can be achieved by statistically sharing channels for adaptive routing of single packet messages with the help of a storage facility at the intermediate nodes. Each packet carries the destination as a part of the header information.

In the traditional message/packet switching, a packet is transmitted out of a node only after the packet is completely received, even if an outgoing channel is free at the intermediate node while the packet is being received. In fact, a packet can be immediately retransmitted after only the reception of the header by

recognizing a free outgoing channel. Thus, message delay can be reduced by buffering only when it is necessary, e.g., when all the channels are busy.

If a packet encounters busy channels at an intermediate node, the outcome is exactly the same as in a traditional packet switched network. On the other hand, if any channel is free, the outcome is similar to a circuit switched system. Thus, the technique used in the present model is an integrated or a hybrid mixture of circuit and packet switching.

The integrated switching scheme is more advantageous than pure circuit or packet switching because it possesses the shared channel properties of packet switching and the dedicated channel properties of circuit switching.

The following assumptions are made for the delay analysis:

- Each message follows Poisson arrival and exponential service statistics, i.e., the arrival times and lengths of individual messages are chosen independently and at random.
- Infinite nodal storage capacity is provided to accommodate unlimited message queueing.
- Deterministic routing for the messages, i.e., a single fixed path between source and destination is assumed.

Let

- $\lambda_i$  = average number of packets per second entering the  $i$ th channel link.
- $1/\mu$  = average packet length (with header) from all the sources (bits/packet)
- $\gamma$  = total arrival rate of packets from all sources
- $C_i$  = channel capacity of  $i$ th channel (bits per sec)
- $C$  = sum of all channel capacities =  $\sum_i C_i$
- $\rho$  = offered load, namely the ratio of average arrival rate of bits into the system to the total capacity =  $\lambda/(\mu C)$
- $d_i$  = average delay for a packet passing through channel  $i$  (transmission plus queueing delay)

When the individual delays at each channel in the system are added, the average time a packet spends in the traditional system is given by [Kleinrock, 1964; Gerla, 1977]

$$d = (1/\gamma) \sum_i \lambda_i d_i \quad (3.17)$$

where

$$d_i = 1/(\mu C_i - \lambda_i) \quad (3.18)$$

Further analysis is required to estimate the savings in queueing and transmission delay for the integrated packet and circuit switching case. For typical packet lengths this savings may not be significant compared to a round trip propagation delay to a geostationary satellite.

The end-to-end system delay can be given by: sum of the satellite processing delay ( $\approx 0.1$  sec), round trip propagation delay ( $\approx 270$  ms), and the queueing and transmission delay.

The protocol described earlier assumes that if the packet is found erroneous and requires retransmissions, the retransmissions are rerouted via a terrestrial path. The overall system delay is considerably reduced in the single hop case, compared to the multiple hop cases. Any time, for a successful packet transmission only one round propagation delay is included. Thus, the protocol of handling retransmission over a terrestrial network is recommended for achieving considerably smaller system delays.

### 3.1.7 Recommendations

The system control issues discussed in the previous subsections do not govern the technology. However, analysis and modeling are necessary for system design guidance and performance evaluation. With the latest available microprocessor and display technology, system control and management functions can be fully realized. Based on the earlier discussions, some of the recommended alternatives include:

- A hybrid control approach including regional and central controls.
- Two assignment schemes, (a) reservation on the message channels and random assignment on control (orderwire) channel, (b) reservation on the message channel and fixed assignment on the control channel.
- A circuit switched, double hop, demand assignment protocol using a central controller.
- A destination variable protocol employing integrated (circuit and packet) switching, single hop, decentralized control, with retransmissions handled by terrestrial network.

To fully evaluate the system performance with reference to control and management aspects, further analysis and simulations should be done in the following:

- Modeling and simulation of the assignment schemes on message/control channels: (a) reservation/random, (b) reservation/fixed.
- Simulations of circuit switched demand assignment (blocked calls cleared/blocked calls delayed) protocol; comparison with the analytical results for given traffic load and blocking probabilities.
- Simulations of destination variable demand assignment protocols and comparison with the analytical results, specifically, analysis of throughput delay improvements by handling retransmissions over the terrestrial network.
- Simulations of on-board controller assignments/scheduling the prioritized traffic (data/voice/video) based upon the packet length and acceptable delay, etc.; comparison of throughput delay performance improvements over ground scheduling.
- Comparison of switching techniques (circuit, packet and integrated) to obtain the performance bounds for a given set of input parameters, e.g., input rate, number of path lengths, and number of channels/link, etc.



### 3.2 MULTIPLE ACCESS/MULTIPLEXING

Principal multiple access (MA) schemes of interest in satellite communications are considered in this section. MA techniques discussed include TDMA, which is most appropriate for non-regenerating, nonlinear satellite repeaters; FDMA, which is best suited to regenerating, (demodulating/remodulating) satellites; and space division multiple access (SDMA), especially multiple beam techniques. Multiple satellites, polarization diversity, and various hybrid MA schemes might also be considered.

Since it is difficult to dissociate methods of on-board switching from MA techniques, the basic kinds of circuit switching and RF switching, e.g., satellite switched TDMA (SS-TDMA), for non-regenerating satellites, and circuit, message, and packet switching for regenerating satellites are reviewed. Message and packet switching ideas require more on-board processing capability.

In addition to explaining and classifying various MA and switching techniques, examples showing typical performances, applicabilities, and implementations are provided. Recommendations are derived from this survey for system considerations: appropriate classes of user terminals; computational complexity; and technological development.

#### 3.2.1 Overview of Techniques

Several known types of MA and kinds of network switching are categorized. The emphasis is on features closely related to on-board satellite processing.

MA techniques can be distinguished according to the following somewhat overlapping categories: time, frequency, space, and code division MA (TDMA, FDMA, SDMA, and CDMA), spread spectrum MA (SSMA), random access (RA), and DAMA. Each of these schemes is more or less suited to particular applications and/or classes of users. Numerical examples are given later in the report to emphasize some differences among these MA techniques. Only qualitative distinctions are provided here.

The wide coverage and broadcast nature of the satellite communications medium are among the most important properties that can be exploited in considering various MA schemes. Not only are widely dispersed users able to gain easy access to satellite capacity but they may employ downlink transmissions to advantage in modifying their own communications protocols. On-board processing

permits the adaptive allocation of multiple antenna beams, improved link performance, internodal message format conversions, and/or crosslinking among satellites, all of which can enhance the flexibility and connectivity of a communications system.

#### 3.2.1.1 Types of Multiple Access

The time-frequency-space MA resources can be split in several ways, as follows:

TDMA. In TDMA, each user is allocated a fixed set of time slots during which the entire capacity of a satellite channel is devoted to that user. The principal advantages are a guarantee of co-channel interference free communications during allotted times and the avoidance of downlink intermodulation interference in the case of non-linear transponder satellites. The primary disadvantages are the potential waste of satellite capacity with lower duty factor users and the possibility of stringent timing and/or ranging requirements that may be difficult to satisfy with lower cost terminals. In addition, the burst rates required by digital TDMA may either exceed the signal processing state-of-the-art or the capability of disadvantaged terminals.

FDMA. FDMA is analogous to TDMA. Each user is allocated a fixed set of frequencies which are devoted to that user for all time. The advantages include lower burst rates, a guarantee of no co-channel interference on allotted frequencies, and the absence of network synchronization requirements. The principal disadvantages are the waste of satellite capacity when there is nothing to transmit and the introduction of intermodulation problems with non-regenerating satellites without network transmitter power control at the terminals. Furthermore, the replication of parallel RF equipment chains on-board the satellite may exceed the capabilities of the current spaceborne state-of-the-art. Also, a given FDMA channel may be vulnerable to RF interference (RFI).

SDMA. SDMA includes multiple beam antennas and polarization diversity permitting added degrees of freedom for frequency reuse. This leads to more complex spacecraft in providing more RF hardware, fast beam switching, isolation circuitry, etc. Multiple satellites, including those in non-geostationary orbits, allow spatial diversity through a variety of terminal pointing arrangements with shaped RF beams. For purposes of the present study, only a single geostationary satellite employing multiple beams is considered.

CDMA and SSMA. CDMA and SSMA are really two different names for the same generic technique encompassing combinations or hybrids of TDMA and FDMA. Timing and RF replication requirements of TDMA and FDMA, respectively, are relaxed at the expense of some digital coding complexity. This usually presents an attractive implementation tradeoff. CDMA and SSMA are more widely applicable than TDMA and more resistant to RFI than FDMA. These hybrid techniques allow greater MA flexibility by permitting a larger pool of users to be assigned to and utilize the same time-frequency-space dimensions of a satellite uplink.

A regenerating satellite provides an excellent opportunity for the joint demodulation of multiple uplink signals by a single receiving platform. This will provide further performance or bandwidth efficiencies in conjunction with uplink coding to reduce transmitter ambiguities and/or to mitigate interchannel interference [Cover, 1972]. Although this relatively new field is likely to receive considerable attention in the future [Wolf, 1979], it is not treated in detail in this report because additional research is required to assess the potential gains and to discover suitable codes for practical implementation.

RA. RA is best suited to simple, low duty factor, low cost terminals that are completely or almost entirely unconstrained in their transmissions. This makes possible over assignment of many terminals to each satellite and makes better use of satellite capacity for these users than fixed MA schemes. On the other hand, throughput efficiency on RA channels may be degraded from that achievable with other access schemes for higher duty factor users.

Random access techniques have little relevance in this study because the terminals are assumed to be relatively complex and typically handle the traffic of high duty factor users, where a premium is placed on high throughput efficiency.

DAMA. A wide spectrum of DAMA techniques can be considered in any particular application [Kota, 1979]. These can range from purely RA schemes with distributed control to various forms of time varying TDMA or FDMA assignments under centralized control. In most situations it is likely that a hybrid control scheme will be most appropriate from both performance and cost points of view. Examples of the DAMA techniques of greatest interest in the present context are covered in the system control section of this report.

### 3.2.1.2 Multiplexing Dualities

Certain dualities among TDM(A) and FDM(A) schemes are explored to further develop contexts for on-board processing. Dual cases are obtained simply by interchanging time and frequency.

Consider a system involving an arbitrary but fixed size time-frequency signal space with an area  $TW$  equal to the time-bandwidth product, where  $T$  and  $W$  are taken as the practical duration and bandwidth of the overall signal set. (See figure 3-8.)

In FDM(A) the  $i$ th user signal occupies a disjoint frequency slot of bandwidth  $W_i < W$  but the entire epoch  $T$ . Adjacent user bands are separated by a certain center frequency spacing, e.g.,  $\Delta W \geq (W_1 + W_2)/2$ ; the corresponding guard bandwidth between frequency channels is  $\Delta W - (W_1 + W_2)/2$ . Guardbands are necessary to protect against interchannel interference, taking into account insufficient spectral roll-off or signal filtering, carrier frequency uncertainties, etc. Holding the overall bandwidth  $W$  fixed, the system designer has the freedom of selecting different frequency slot assignments from epoch to epoch. This may alleviate intersystem interference to certain users from narrowband RFI sources, for example.

In TDM(A) the  $i$ th user signal occupies a disjoint time slot of duration  $T_i < T$  but the entire bandwidth  $W$ . Adjacent user slots are separated by a certain center slot spacing, e.g.,  $\Delta T \geq (T_1 + T_2)/2$ ; the corresponding guard duration between time channels is  $\Delta T - (T_1 + T_2)/2$ . Guard times are also necessary to protect against interchannel interference, but insufficient timing control or signal gating accuracies, reference timing uncertainties, etc. must be considered. Holding the overall frame duration  $T$  fixed, the system designer may select different time slot assignments from epoch to epoch. This may alleviate intersystem interference to certain users from periodic bursty noise sources, for example.

In the context of satellite communications, one often refers to the signal space organizations of figure 3-8 as FDMA and TDMA when a number of distinct uplink signals are transmitted to a satellite from separate user terminals. Note that any of these signals could still have a sub-signal space organization like those of the figure 3-8, i.e., sub-channels that are either FDM or TDM, where the multiplexing at this next lower level is accomplished by the corresponding terminal.

Several kinds of on-board processing can be realized in the satellite. One of the simplest forms is frequency translation between each uplink and downlink pair on a single channel per

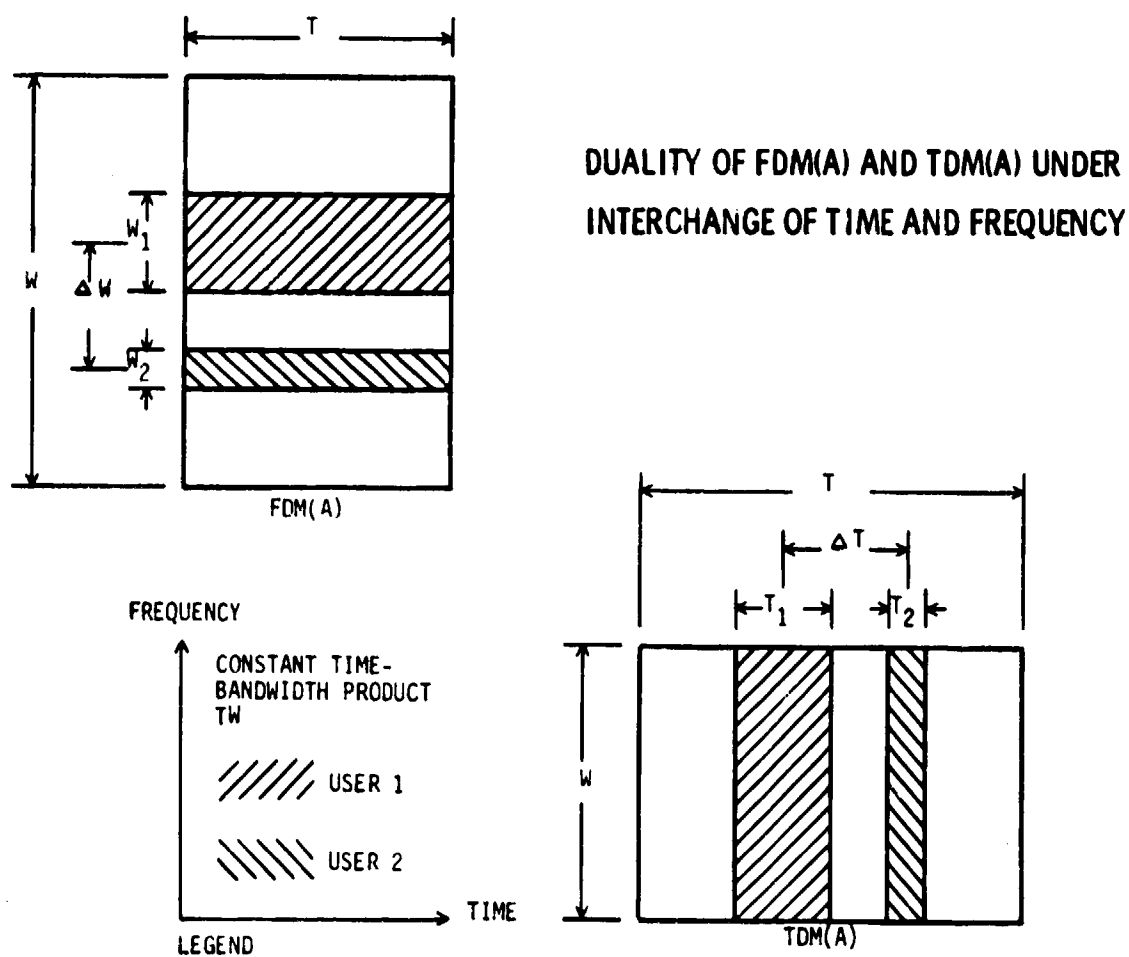


Figure 3-8. Duality of FDM(A) and TDM(A) Under Interchange of Time and Frequency

carrier (SCPC) basis. In the case of FDMA, the frequency bands can be shuffled for the downlink, if desired, because a separate satellite transponder must be devoted to each frequency channel. Frequency translation helps mitigate frequency interference problems with hardware implementations of uplink and downlink interfaces on-board a satellite and permits the use of a common antenna. Hence, separate uplink and downlink frequency bands are usually allocated for satellite communications, e.g., nominally 30 GHz uplink and 20 GHz downlink. An exception is the aeronautical and maritime mobile-satellite band of 43 to 48 GHz where particular uplink and downlink frequency bands are not specified.

Since the dual of frequency translation is time translation, it is conceivable that the uplinks and downlinks could be separated in time instead of frequency to completely avoid frequency interference problems across the uplink and downlink interfaces. This would mean that the uplink and downlink frequency bands could be identical and would permit some of the frequency allocation burden to be converted to one of time slot allocations for uplinks and downlinks. Time translation would require means for delaying uplink signals prior to downlink retransmission. This could be accomplished with analog or digital delay lines depending on whether the satellite is the regenerative type. In the case of TDMA, the time slots of figure 3-b could then be shuffled on-board the satellite, if desired.

It is recommended that greater attention be given to time translation as an alternative to frequency translation in future satellite systems. The possible benefits of buffering and/or storing digital data in high capacity low cost memories\* on-board the spacecraft seem worthy of further investigation. The advantages of such an approach may be especially attractive for non-real time data communications where significant additions to round-trip propagation delays for a geostationary satellite are relatively unimportant.

A third dimension, that of geometry or physical space, can be added to the signal space of figure 3-8. In this report the third dimension is embodied by the concept of multiple beam spacecraft antennas for frequency reuse. Thus, an additional layer of satellite processing techniques can be applied to directing a mix of

---

\* The non-volatile radiation resistant bubble memories under intensive current development [Mier, et al., 1977] are prime candidates for spaceborne applications.

FDMA or TDMA traffic in a given uplink beam to the appropriate downlink beams. In a regenerating satellite, this process may include a reorganization of data streams and modulation formats as well. For example, a set of FDMA continuous phase frequency shift keyed (FSK) signals in one uplink scanning beam might be transformed into portions of TDM QPSK signals in one or more fixed beam downlinks. Such on-board processing capabilities permit the system designer to attain the flexibility and performance advantages that can accrue by completely separating the uplinks and downlinks. In particular, accommodations can be made for smaller terminals and difficult satellite design problems.

#### 3.2.1.3 Types of Switching

Circuit switching is by far the most commonly understood form of switching and the technique usually associated with satellites as switchboards-in-the-sky [Fordyce, Jaffe, 1978]. Nevertheless, message and packet switching offer greater potential in reaping the benefits of the MA/broadcast attributes of the satellite medium for many possible users. Again, in this introductory subsection, only the qualitative distinctions among switching techniques are discussed. Numerical examples follow later in the report.

Circuit. This form of switching is most appropriate for higher duty factor users where set-up time is short compared to the duration of the call. Minimal on-board memory is required for regenerative repeaters of digital direct-to-user traffic.

Message. Message switching of digital traffic is often thought of in the context of store-and-forward networks where the entire message is stored at the satellite before retransmission. This can be a more attractive technique for lower duty factor users that require path diversity for improved reliability and/or that may tolerate longer end-to-end message delays.

Packet. This is a special case of message switching where there are one-packet messages. In general, less on-board storage is required, and there is more path flexibility which may lead to greater message reliability and/or shorter message delays. On the other hand, more severe network stability problems due to individual link overloading may arise compared to message switching.

In this report the term packet switching may also refer to an on-board processing technique where only header information is interpreted and stored in the satellite while the immediately

following message portions of uplink packets are routed to the appropriate downlink with minimum buffering. Since the satellite does not necessarily wait for the entire message to be received, this form of packet switching is not the traditional one in the store-and-forward sense.

#### 3.2.1.4 Types of Packet Transmission

Four common types of packet switching multiple access are reviewed briefly below.

Pure and Slotted ALOHA. In Pure ALOHA, each terminal transmits a packet of information essentially at random without regard to any network control discipline. Theoretically, any set of user packets which overlap in the time-frequency-space resource of the channel are lost. For the time domain, the maximum channel throughput is approximately 18%.

By defining time slots equal to a packet length and requiring any packet transmission to fill a slot, the throughput can be doubled to about 37% in Slotted ALOHA. However, the slots for packet transmission are still selected at random.

Capture ALOHA. A downlink limited processing satellite can spend uplinks and/or on-board memory to improve channel throughput over that of Slotted ALOHA. This can be accomplished by effectively capturing one packet for downlink transmission among any set of colliding packets. Unfortunately, in addition to requiring more satellite complexity, the delay performance may degrade compared to a single access link.

Carrier Sense MA. If packet lengths are significantly longer than round trip propagation delays, then carrier sense MA (CSMA) techniques can be applied to advantage, even in satellite links. In this case, users refrain from transmitting a packet in a slot where a higher priority user is transmitting. Generally, this technique will only apply to non-real time data transmissions from smaller numbers of users.

The applicability of any MA/switching scheme is sensitive to the number of concurrent satellite accesses and the channel throughput or network operating point. It is impossible to declare one technique uniformly superior to any other without regard to desired user requirements and equipment complexities.



### 3.2.2 Multiple Beams

Some MA/switching techniques are explained by way of example in this subsection. Although emphasis is placed on multiple beams for real time point-to-point data applications, most of the results can be applied to other types of service as well.

The primary motivations for multiple satellite beams are to provide for frequency reuse and coverage flexibility. Alternative frequency plans for various beam patterns are not considered here. It is assumed that beam footprints are sufficiently separated in frequency or time so that the frequency reuse factor is identical to the number of simultaneous beams. Multiple beam organizations are explored in addressing questions of throughput; total number of users possible; and switch sizes, processing speed, and memory requirements on-board a satellite.

Secondary motivations for multiple beams are obtained indirectly by the desire for a high satellite antenna gain. This permits terminals to have lower effective isotropic radiated power (EIRP) and gain to noise temperature ratio (G/T), and is consistent with the trend toward many lower cost terminals and larger, more complex platforms in space. Higher link margins, improved system availabilities in the presence of rain and other sources of atmospheric attenuation, and less intersystem interference result from larger satellite antennas. Since higher gain typically means smaller coverage areas, multiple beams are implied.

A lower satellite noise temperature contributes in direct proportion to lower terminal EIRP only as long as receiver noise dominates other forms of uplink interference. Higher satellite RF power also contributes directly in allowing lower terminal G/T. However, one generally expects a significant fraction of on-board power to be devoted to functions other than downlink transmission. Thus, higher satellite antenna gain is the most direct way of providing for lower cost terminals [Morrow, 1976].

Fixed beams simplify terminal operation and are natural for regions of high expected traffic intensity, such as metropolitan areas. However, multiple fixed beams translate into a number of parallel chains of RF processing hardware which increase in direct proportion to the number of fixed beams. This may imply too much on-board RF complexity and suggests the possibility of scanning beams, though the complexity reduction factor may be less than the number of fixed beams that a scanning beam replaces.

It may be possible to share a single reflector among fixed and scanning beams using multiple feeds and polarization diversity [Reudink, 1976]. The difficulty of constructing large spaceborne phased arrays suggests that only a handful of simultaneous scanning beams is feasible. These issues require further study and are beyond the scope of this report.

Although scanning beams would seem to save RF hardware and simplify downlink intermodulation problems, continued technological advances may make scanning beams unnecessary. Power-hungry TWTs may be replaced by smaller, more efficient, solid state devices that can be switched essentially instantaneously. Therefore, one may not need to switch one TWT among several different beams to save on-board power. Intermodulation can be controlled simply by more physical separation between 20 GHz horns, for example, as larger structures in space remove present spacecraft size constraints [Edelson, Morgan, 1977].

A principal argument for scanning beams is that they provide one possible way of handling varying traffic demands in a flexible manner if satellite capacity is not saturated by the actual traffic in the fixed beams. One can envision pointing a scanning beam to areas not covered by a fixed beam to gather additional traffic on a demand assignment basis or even to fixed beam areas where demand exceeds the fixed beam capacity. In the latter case, a disjoint frequency band for the scanning beam would be required.

If there are scanning beams, they would seem to be more natural for regions of relatively low duty factor traffic, such as rural areas. As long as these beams do not scan too fast, they should not unduly increase the cost of the associated terminals. Fewer and larger fixed beams could service the remote areas, but this would deprive the smaller terminal of the higher satellite antenna gain advantage.

#### 3.2.2.1 Assumptions

It is assumed that each of four sets of beams, viz., fixed uplink, scanning uplink, fixed downlink, and scanning downlink, is comprised of disjoint but otherwise indistinguishable beams. Four distinct groups of users operating at the standard transmit (TX) carrier data rates (T1: 1.544 Mb/s, T2: 6.312 Mb/s, T3: 44.736 Mb/s, T4: 276.176 Mb/s) are assumed. Equal duration slots and equally spaced carrier frequencies will always be assumed when referring to any TDMA or TDM and FDMA or FDM multiple access or multiplexing schemes. These assumptions are made for convenience in computing

average parameter values for predicting expected behavior and sizing the system. Such constraints would be overly restrictive and unnecessary in a detailed design.

The focus of discussion is on a single geostationary domestic satellite centered at an available mid-United States longitudinal parking orbit location. The round-trip propagation delay between any pair of user terminals is approximately  $250 \pm 20$  msec. Furthermore, each terminal is presumed to be full duplex and to have only one antenna for both transmitting and receiving in the allocated uplink and downlink frequency bands at 30/20 GHz. A more complex system could employ some combination of multiple satellites, terminal antennas, and frequency bands, but such concepts are not considered here.

TDMA and FDMA are taken as the generic possibilities for multiple access in the uplink beams. Similarly, TDM and FDM are the basic multiplexing techniques for the downlink beams. Although this implies 16 possible combinations of multiple access and downlink multiplexing for the four sets of beams, not all these possibilities can be considered in this report. (See the section on architectural alternatives.)

The multiple access schemes and downlink multiplexing techniques assumed are TDMA and TDM for fixed uplink and downlink beams, and FDMA and FDM for scanning uplink and downlink beams. This is consistent with the notion of more complex, higher cost, heavier traffic terminals in metropolitan areas that can handle higher burst rates and more stringent synchronization requirements, and simpler, lower cost, and lower duty factor terminals in remote areas.

A principal thesis of this study is that the smaller the size of user served by direct satellite communications, the greater the potential gains in flexibility and performance from on-board regenerative processing, i.e., demodulating, reformatting, rerouting, and remodulation. One would not expect to perform much demodulation of the highest data rate T3 and T4 carriers of trunking traffic (44.736 Mb/s and 276.176 Mb/s). This is because of current technological infeasibility, and also because there is a pressing need for improved flexibility or performance in such long haul applications. Existing "bent pipe" satellite concepts can be extended in channel capacity and given adequate flexibility through the use of multiple beams and on-board RF switching techniques. Current efforts in the higher frequency bands, multiple beams, and

RF switching will also help to alleviate the spectral and orbital crowding problems that continue to intensify as capacity demands increase.

On the other hand, for burst rates in the hundreds of Mbps range that push the state-of-the-art, more spacecraft power efficiency may be obtained with on-board demodulation (using surface acoustic wave (SAW) devices?) and no on-board switching. The switching function may be better performed by frame organization and network synchronization at the terminals. The satellite would then be used as a very efficient regenerative "bent pipe". This tradeoff deserves further analysis.

At the other end of the user spectrum, expansion of the public service market for very small inexpensive terminals depends on affordable technology, as in the case of hand-held calculators. Vigorous pursuit of the necessary high risk satellite technology is not evident [Bekey, 1978].

The T1 and T2 carrier (1.544 Mb/s and 6.312 Mb/s) classes of direct-to-satellite user which should provide link data rates representative of the maximum rate feasible for on-board demodulation are emphasized here. The data traffic is presumed to have a real time ingredient which will introduce processing speed and buffering constraints into any system design. In addition, although a standard T1 carrier consists of twenty-four 64 kb/s telephone channels, for example, processing individual telephone conversations at the satellite is not discussed. The TX carriers are considered on a strictly point-to-point basis. There is no attempt to handle teleconferencing, at least initially. For example, the notion of relaying information contained in an elementary uplink transmission to several downlink destinations is not treated in this section. Although this simplifies the on-board control, reformatting, and rerouting procedures, one should recognize that the inherent broadcast nature of the satellite medium is not used to full advantage with this example.

The modulation format for bandwidth efficiency is not a critical issue with coordinated uplink users whose signals are demodulated on-board the spacecraft. Groups of synchronized QPSK signals with a nominal 1 b/s/Hz (or 1 bit/cycle) can be envisioned without much loss in generality. With uncoordinated uplink signaling (different burst rates, unsynchronized timing, no power or frequency control, etc.), MSK-type signaling may be superior for bandwidth conservation. (See section 3.3.)

### 3.2.2.2 Notation

The following variables are defined in relation to the four sets of beams and the four groups of users. Here user refers to a terminal that is directly linked with the satellite. Variables may be subscripted with f for fixed, s for scanning, u for uplink, d for downlink, and i = 1,2,3,4 for T<sub>i</sub> carrier.

A	number of disjoint coverage areas
B	user burst rate
b	user bits per frame; also bits per burst
$\beta$	bandwidth efficiency factor in b/s per Hz or bits per cycle (single channel per carrier packing factor)
D	on-board processing delay
F	frame duration
L	mini-packet length (bits)
M	number of simultaneous beams (frequency reuse factor)
m	number of simultaneous users or number of user signals present at a given instant
N	number of separate (FDMA) carriers per beam
n	number of concurrent users or total number of users serviced at a specified data rate
Q	memory size (bits)
R	user data rate
r	burst to data rate ratio; also, frame to slot or dwell ratio
S	maximum switching rate
s	number of mini-slots per slot (relates mini-packet length to bits per burst)
T	slot duration of fixed beam or dwell time of scanning beam; also, burst duration; also, mini-slot duration (with subscript m)
t	throughput
W	instantaneous bandwidth

### 3.2.2.3 Parametric Relationships

Parametric relationships are identified and tradeoffs among various systems parameters are illustrated in this section. First, the fixed and scanning beam services are treated independently to show how the number of users and throughput depends on data rate for a given number of beams and a certain number of carriers per beam. Then, a discussion of frame structure shows the relationship between packet lengths, slot durations, bits per burst, and frame intervals. Finally, processing delay is examined for real time data communication as a function of burst rate disparity between the uplink and downlink.

Throughputs for Fixed and Scanning Beam Services. Under the point-to-point assumption, the number of users and the throughput for the fixed and scanning beams can be computed for either the uplinks or the downlinks. Since a balance must be achieved on the average, the results for the uplink calculation must equal those for the downlinks.

First consider a set of  $M_f$  fixed beams on either the uplink or the downlink. If there are  $N_f$  carriers per beam, and if the average user burst to data rate ratio is  $r_f = B_f/R_f$ , then there are  $m_f = M_f N_f$  simultaneous fixed beam user signals and  $n_f = m_f r_f$  concurrent\* fixed beam users. If the instantaneous signal bandwidth in a fixed beam is defined as  $W_f = N_f B_f / \beta_f$ , when  $\beta_f$  is the SCPC packing factor, then the fixed beam throughput is given by

$$t_f = n_f R_f = m_f B_f = N_f N_f B_f = \beta_f N_f W_f. \quad (3.19)$$

These relationships are displayed parametrically in figure 3-9 for a burst rate of 600 Mb/s, which is taken as the nominal maximum feasible value for the system envisioned. This is based on digital, quadrature carrier modulation (e.g., QPSK) and the availability of off-the-shelf equipment able to operate at 300 MHz clocking rates by the time the system is implemented [Acampora, Langseth, 1979]. For example, with  $N_f = 2$ , the figure shows a throughput of 24 Gb/s from 40 simultaneous or 536 concurrent T3 users in 20 fixed beams. The bandwidth of each of the two carriers is taken as 600 MHz, the double-sided zero-crossing bandwidth of a QPSK modulation spectrum. With a carrier packing factor of  $\beta_f = 1$  for synchronized user signals, the total fixed beam instantaneous bandwidth is  $W_f = 1.2$  GHz. Note that the curves are useful for other burst rates by simply scaling the number of users and throughput by the difference in burst rates.

\*In this case, the number of users being serviced at the data rate  $R_f$ ; different sets of  $m_f$  user signals appear at the burst rate  $B_f$  during each of  $r_f = B_f/R_f$  TDM slots. (See Generic Frame Structures, page 72.)

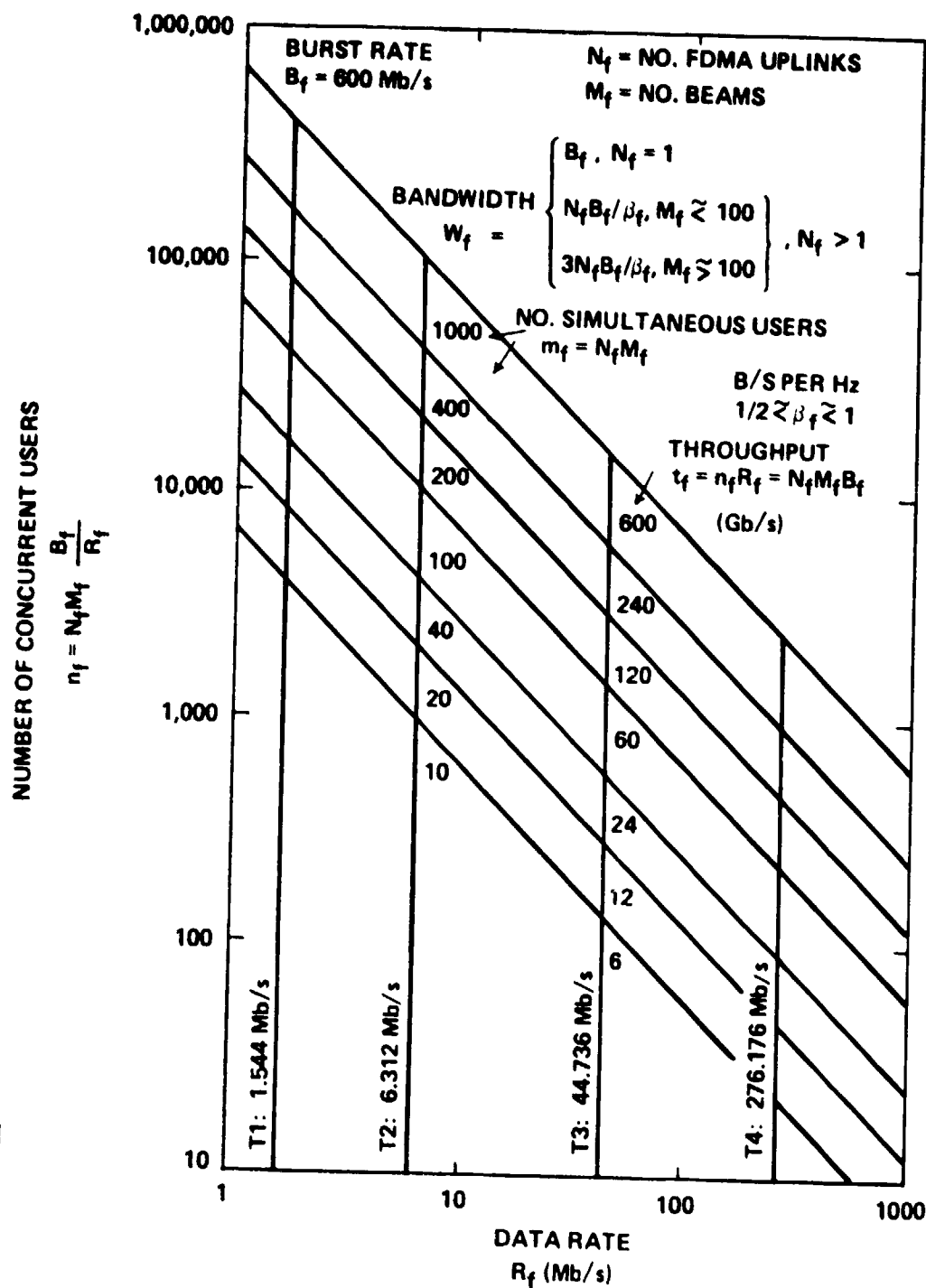


Figure 3-9. Throughput for Fixed Beam Service

When the fixed beam coverage areas become too close or overlap, it may be necessary to utilize a frequency plan where the full frequency reuse factor, equal to the number of simultaneous beams  $M_f$ , is not attained. In figure 3-9, a three frequency plan has been assumed when  $M_f$  exceeds approximately 100, corresponding roughly to the number of distinct coverage areas envisioned for CONUS. This would reduce the frequency reuse factor to  $M_f/3$ .

The case for scanning beam service is quite similar. The throughput

$$t_s = n_s k_s = m_s B_s = N_s N_s B_s = \beta_s M_s W_s \quad (3.20)$$

for the scanning beams is shown graphically in figure 3-10 in the same fashion as for fixed beams except that the burst rate has been reduced by a factor of 10 in deference to the smaller T1 and T2 users of the scanning beams. In this case, 40 simultaneous or 1554 concurrent T1 users in four scanning beams yield a throughput of 2.4 Gb/s, for example. With ten FDMA carriers, each with a double-sided zero-crossing bandwidth of 60 MHz, and a packing factor of  $\beta_s = 1/2$  for unsynchronized user signals, say, the total scanning beam instantaneous bandwidth is  $W_s = 1.2$  GHz. Again, note that a different bandwidth formula would be used in figure 3-10 in the unlikely event that the number of scanning beams approaches the number of distinct coverage areas, nominally 100 for CONUS.

Generic Frame Structures. For the multiple access scenarios assumed for the fixed and scanning beams, no user transmits or receives meaningful data on a continuous basis. Users in a fixed beam operate in a time division mode, while users in a scanning beam can operate only when within the dwell of the beam. In other words, each user transceives data in bursts at a burst rate  $r$  times higher than the data rate, where

$$r = b/R = F/T \quad (3.21)$$

is the burst to data rate ratio or the frame to slot or dwell time ratio.

A generic frame structure is shown in figure 3-11. It is assumed that each user operates in an equal duration portion of each frame, called a slot. This slot corresponds to a TDM(A) slot for a fixed beam and a dwell time of a scanning beam. In the scanning beam case, it is assumed that the users operate in an FDM(A) mode while in the beam. Each slot is composed of subslots which can be of variable length because users may have different amounts of data to communicate among different locations but at the same burst rate.



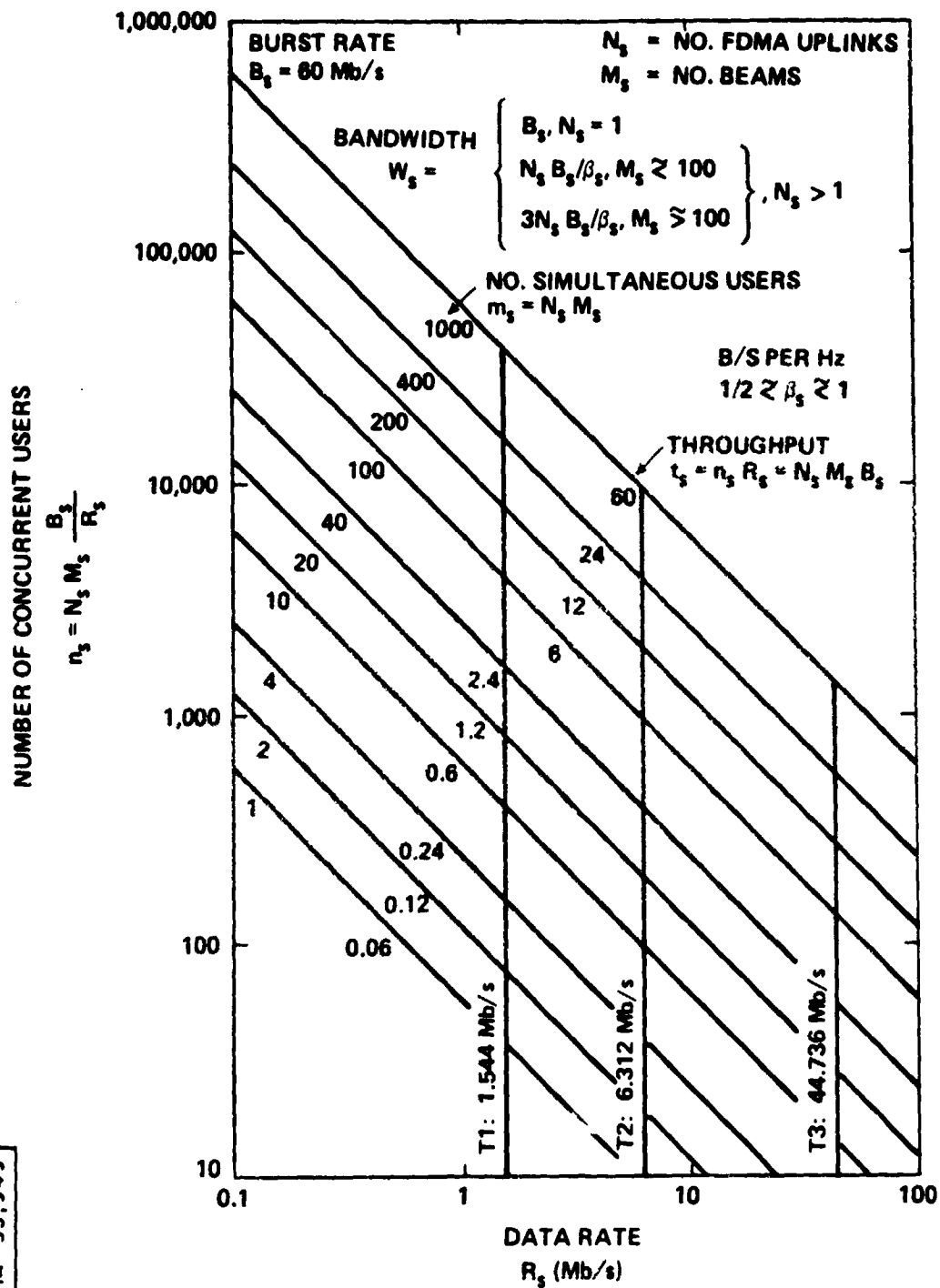


Figure 3-10. Throughput for Scanning Beam Service

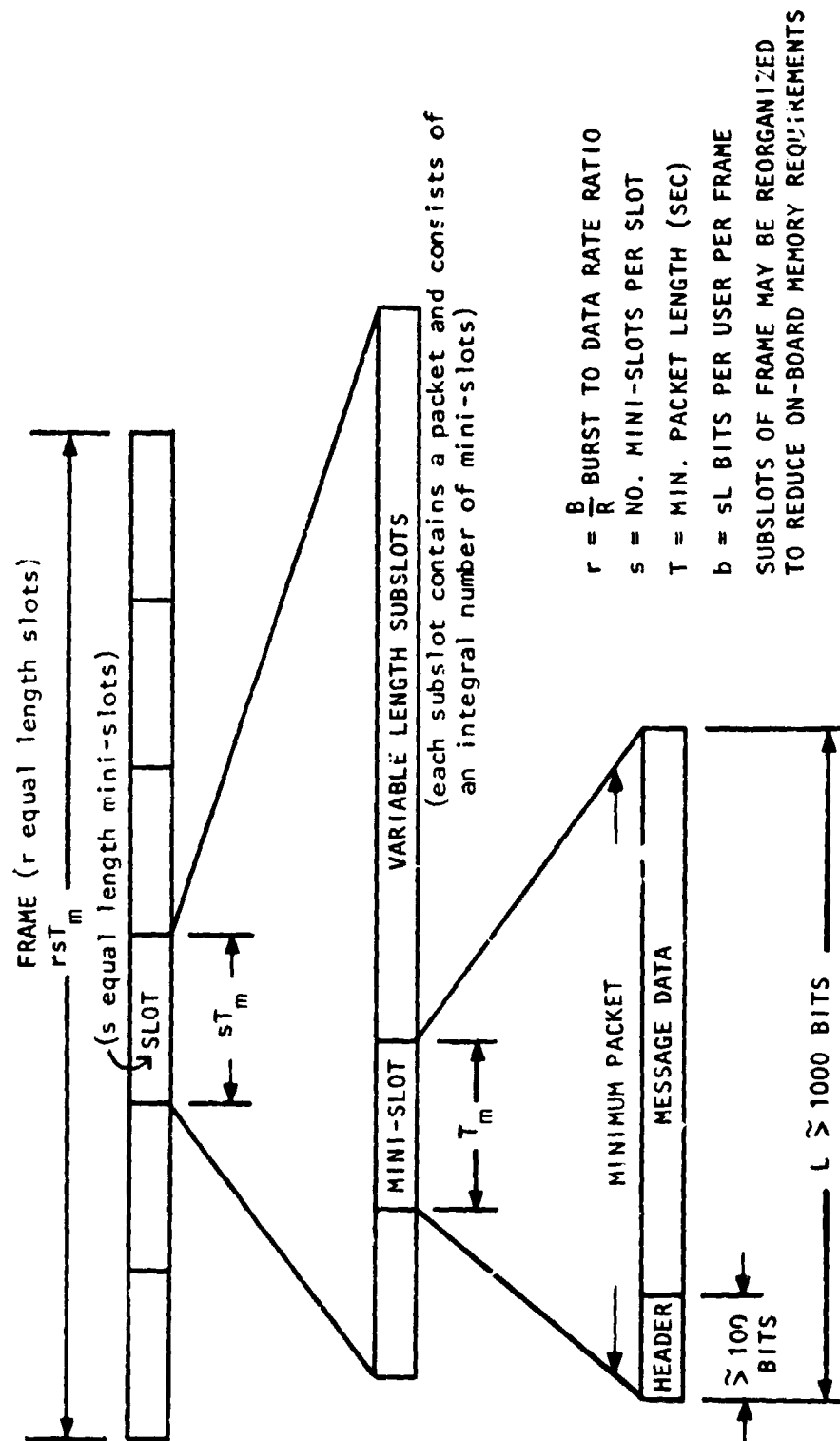


Figure 3-11. Generic Frame Structure

In an actual implementation, the subslots would be rearranged within the frame to accommodate multiple beam organizations of traffic flow which require no on-board memory in the case of RF switching and less memory for regenerated data. Thus, although each slot of figure 3-11 may be segmented and dispersed throughout the frame, this is irrelevant for the nominal calculations of this subsection.

For convenience of design and implementation, each subslot is composed of a variable but integral number of mini-slots of duration  $T_m$ . Let there be  $s$  mini-slots per slot, i.e.,  $T = sT_m$ ; the integer  $s$  may range from one to a hundred, or so. The mini-slot interval determines the maximum switching rate  $S$  for changing connections among users, i.e.,

$$S = 1/T_m = s/T. \quad (3.22)$$

From the standpoint of overhead inefficiencies, there is a lower limit that can be placed on the mini-slot length. In packet switching, for example, header information may require up to about 100 bits, and packets should be at least 1000 bits long to maintain a 90% packet efficiency of message to total data. If  $L = BT_m$  is the minimum number of bits in a packet, then the number of bits per user per frame is

$$b = sL = sBT_m = BT. \quad (3.23)$$

The frame duration is  $F = rT = BT/k = b/R$ , from equations (3.21) and (3.23). Design curves for obtaining  $b$  as a function of  $R$  for various frame intervals are provided in figures 3-12 and 3-13 for the fixed and scanning beam services.

Alternatively, knowing the user data rate  $R$  and the bits per user per frame  $b = sL$ , or the maximum length user packet assuming a continuous burst occupying a contiguous slot of a frame, the frame duration is easily obtained from the design curves of figure 3-14. For example, packet lengths of 16,776 bits for 13 users imply a frame of 375  $\mu$ sec.

For slots of 125  $\mu$ sec duration, e.g., only three T3 users per frame are handled, this implies a burst rate of 134 Mb/s. For a maximum  $s$  of 100, the worst case switching rate would be  $S = 800$  kHz, from equation (3.22). However, with a maximum burst rate of 600 Mb/s, at most thirteen T3 users could be accommodated. For the same frame interval and packet length, this implies a slot duration of 28.8  $\mu$ sec and a maximum switching rate of  $S = 3.5$  MHz. Switching rates such as these should not be a problem for either RF or digital

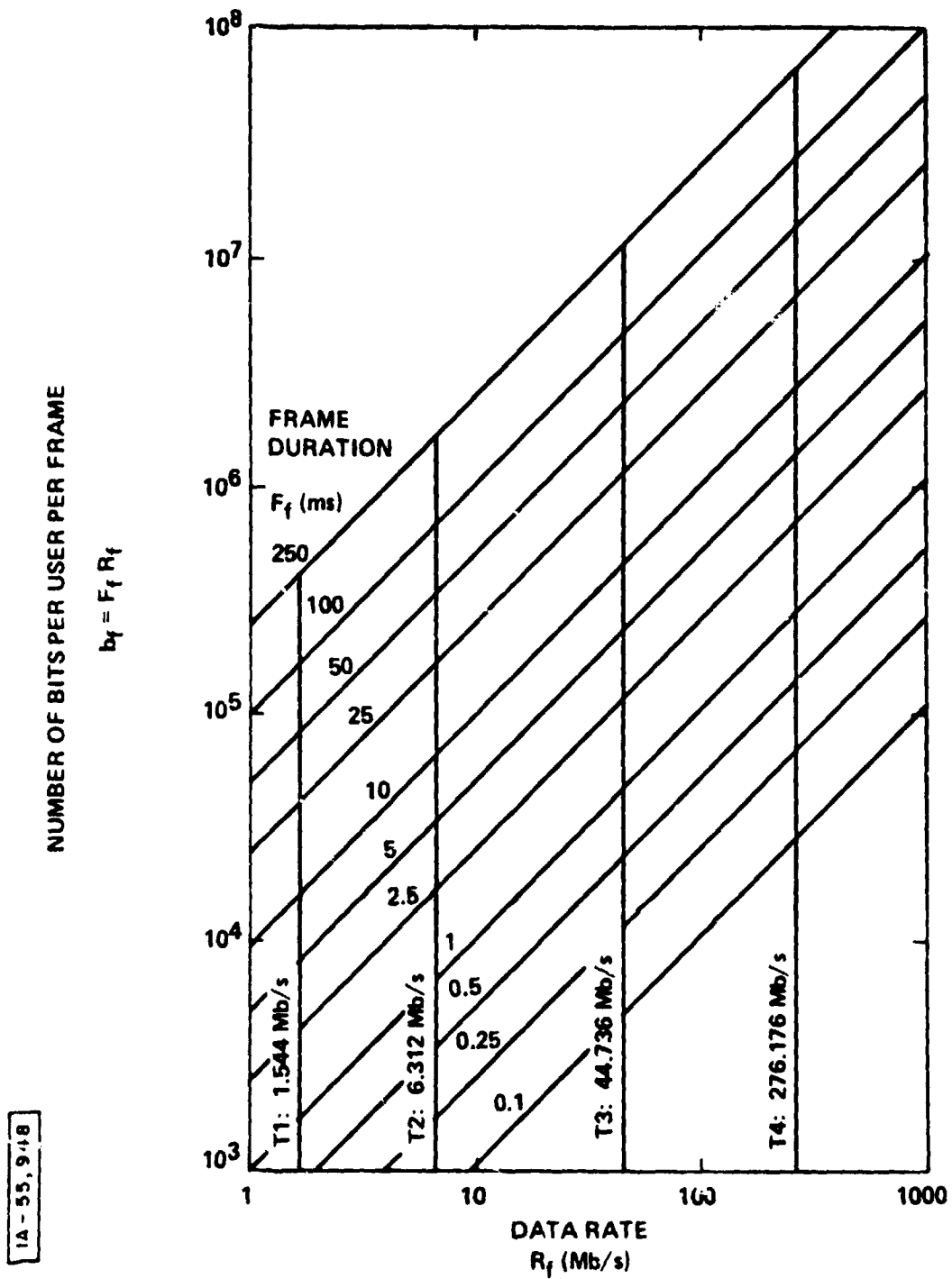


Figure 3-12. Burst Sizes for Fixed Beam Service

14-55,947

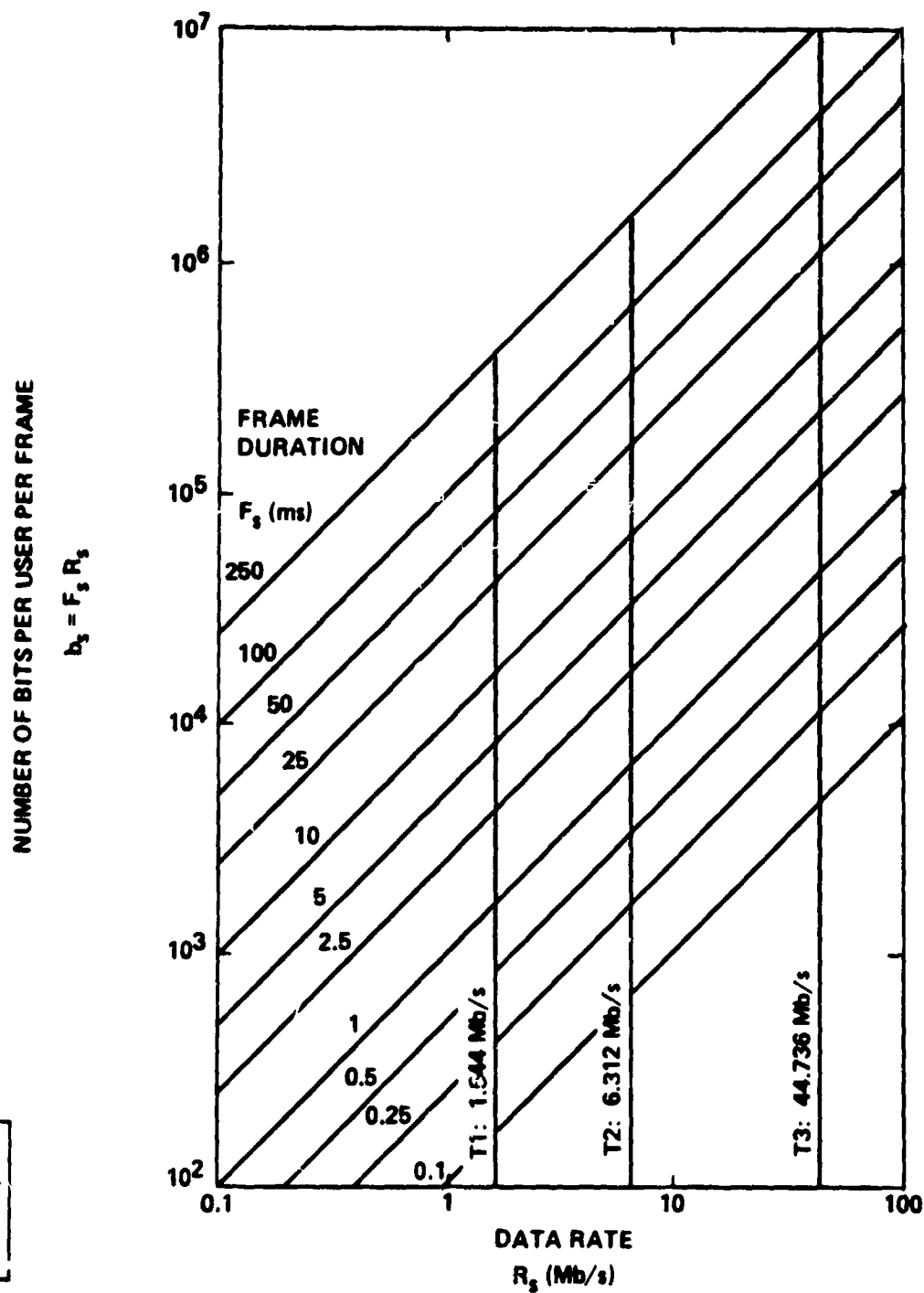


Figure 3-13. Burst Sizes for Scanning Beam Service

1A-55,946

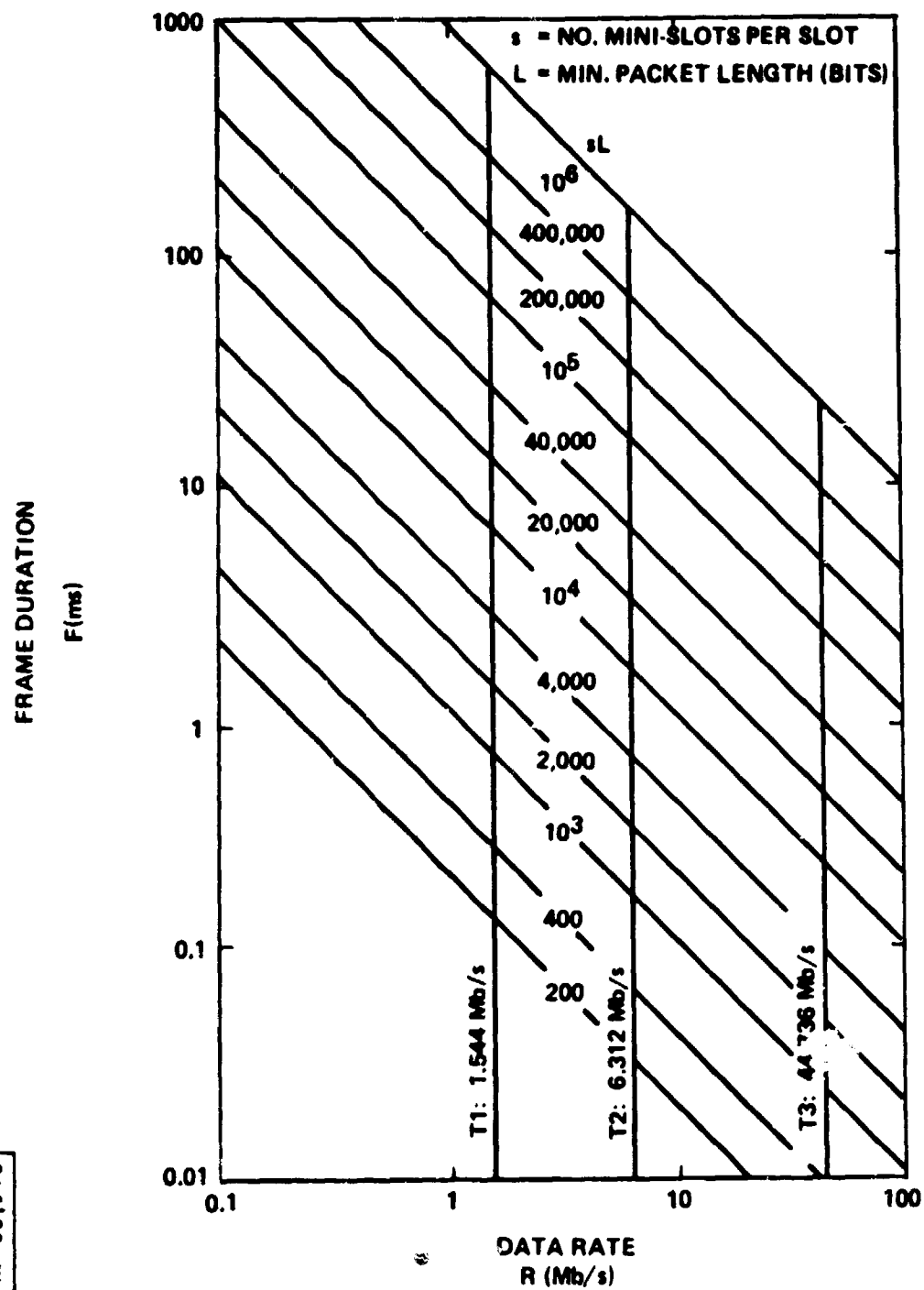


Figure 3-14. Frame Lengths for Either Service

switching, even though switching speeds must be at least an order of magnitude faster than switching rates for sake of efficiency.

On-Board Processing Delays. With on-board RF switching of the T3 and T4 trunking traffic, the uplink and downlink burst rates are identical. However, with on-board regeneration of the T1 and T2 direct to user traffic, there is the possibility of burst rate disparity between an uplink and the corresponding downlink because of changes in data or multiplexing formats, conversions between fixed and scanning beams, etc. For instance, data from a large TDMA user in a fixed beam might be demodulated at a high burst rate on-board the satellite and retransmitted at a lower burst rate in an FDMA downlink format to a smaller user in a scanning beam. On the other hand, FDMA uplink traffic at lower burst rates in a scanning beam might be regenerated in a higher burst rate TDM format for downlink transmission in a fixed beam. In any event, additional processing delays can be introduced in a regenerative repeater when there is an uplink and downlink burst rate disparity. It is important to estimate this processing delay, especially in real time data communications, to insure that performance will not be seriously affected.

Refer to figure 3-15 for the case where the uplink burst rate  $B_u$  exceeds the downlink burst rate  $B_d$ ; the processing delay is seen to critically depend on the burst rate ratio  $B_u/B_d$  and the downlink frame duration  $F_d$  as

$$D = (B_u/b_d - 1)F_d + T_u, \quad B_u > B_d \quad (3.24)$$

where  $T_u$  is the uplink burst duration. This formula follows from the double buffer mechanism indicated, where  $b_u$  bits are accumulated on the uplink in the time it takes one word of only  $b_u B_d/B_u$  bits to be retransmitted on the downlink. The next word cannot be retransmitted until the next downlink time slot for the same destination becomes available, an interval equal to the downlink frame duration, on the average. If the burst rates are identical, the processing delay reduces to only the burst duration, which is then the same for the uplink and downlink, i.e., if  $B_u = B_d$  in equation (3.24),  $D = T_u = T_d$ , the usual processing delay. Note that, in general, the downlink burst of net duration  $T_d$  (in non-real time) is separated by several sub-bursts, each of duration  $T_u$ , where the separation between sub-bursts is  $F_d$ .

The analogous case when the downlink burst rate exceeds the uplink burst rate is depicted in figure 3-16. The processing delay is then given by

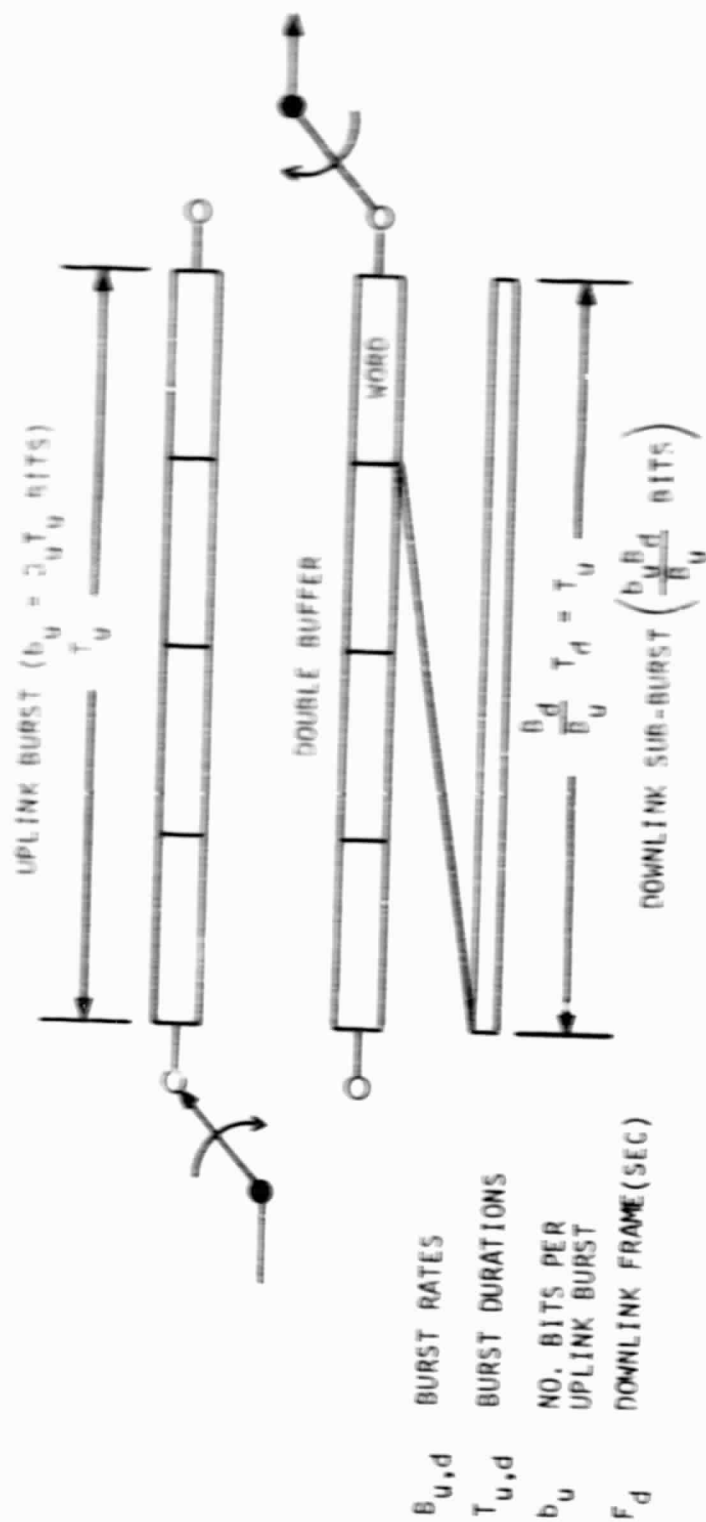


Figure 3-15. Processing Uplink Burst as Separated Downlink Sub-Bursts ( $B_u > B_d$ )



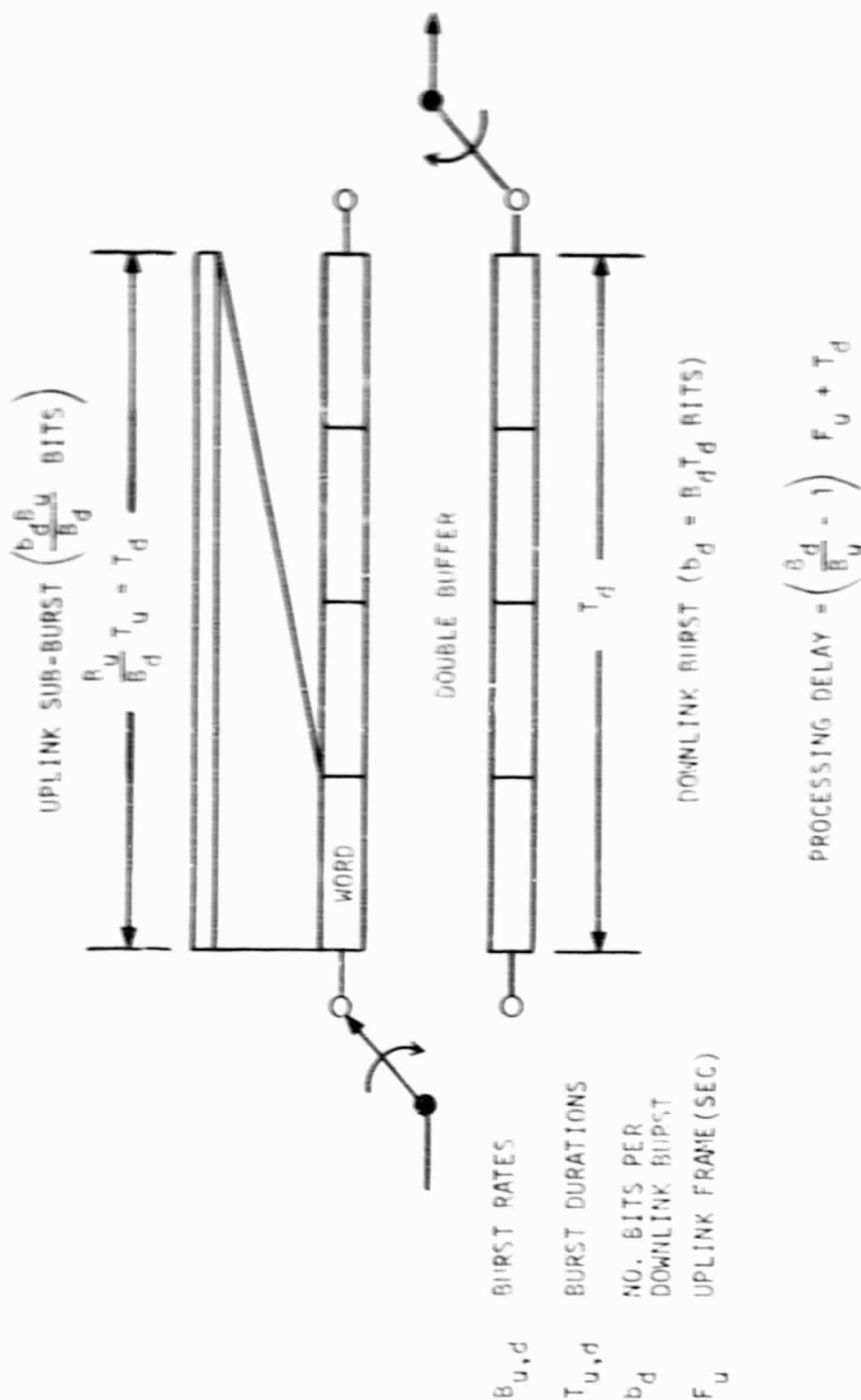


Figure 3-16. Processing Downlink Burst as Separated Uplink Sub-Bursts ( $b_u < b_d$ )

$$D = (B_d/B_u - 1)F_u + T_d, \quad B_u < B_d. \quad (3.25)$$

here the downlink buffer can empty completely in the time necessary to accumulate just one uplink word of  $b_d B_u / b_d$  bits. The next word in the uplink buffer cannot be accumulated until the next uplink time slot from the same source, an interval equal to the uplink frame duration, on the average.

Since the buffers in these examples fill and empty at different rates, some detailed memory organization is required to avoid unnecessary buffer overflows or unused capacity. On the average over time and many uplinks and downlinks, it should be possible to design efficient buffer/memory processing algorithms for the regenerated signals. It is beyond the scope of this report to do so.

The main point is that it is necessary to limit the uplink and downlink burst rate disparity to avoid adding significantly to the propagation delay already experienced with a geostationary satellite. Since the one-way propagation delay to geostationary altitude is approximately 1/8 sec, one would like the on-board processing delay to be no more than about 12.5 msec for real time data communications. Figure 3-17 shows the largest permissible burst rate disparity for a given frame interval. For example, for a 1.25 msec uplink frame, the downlink burst rate can be no more than ten times the uplink burst rate for an acceptable real time processing delay.

Other System Considerations. The total number of simultaneous uplink and downlink beams is

$$M_u = N_{fu} + N_{su} \quad (3.26a)$$

and

$$M_d = N_{fd} + N_{sd}. \quad (3.26b)$$

With no beam overlaps these numbers are equivalent to frequency reuse factors.  $M_d$  may be limited to ten or twenty if downlink intermodulation interference is a problem. However, it is conceivable that more simultaneous downlink beams can be engineered for a very large spacecraft.

The number of disjoint uplink and downlink coverage areas is

$$A_u = N_{fu} + N_{su} F_{su} / T_{su} \quad (3.27a)$$

1A-55,945

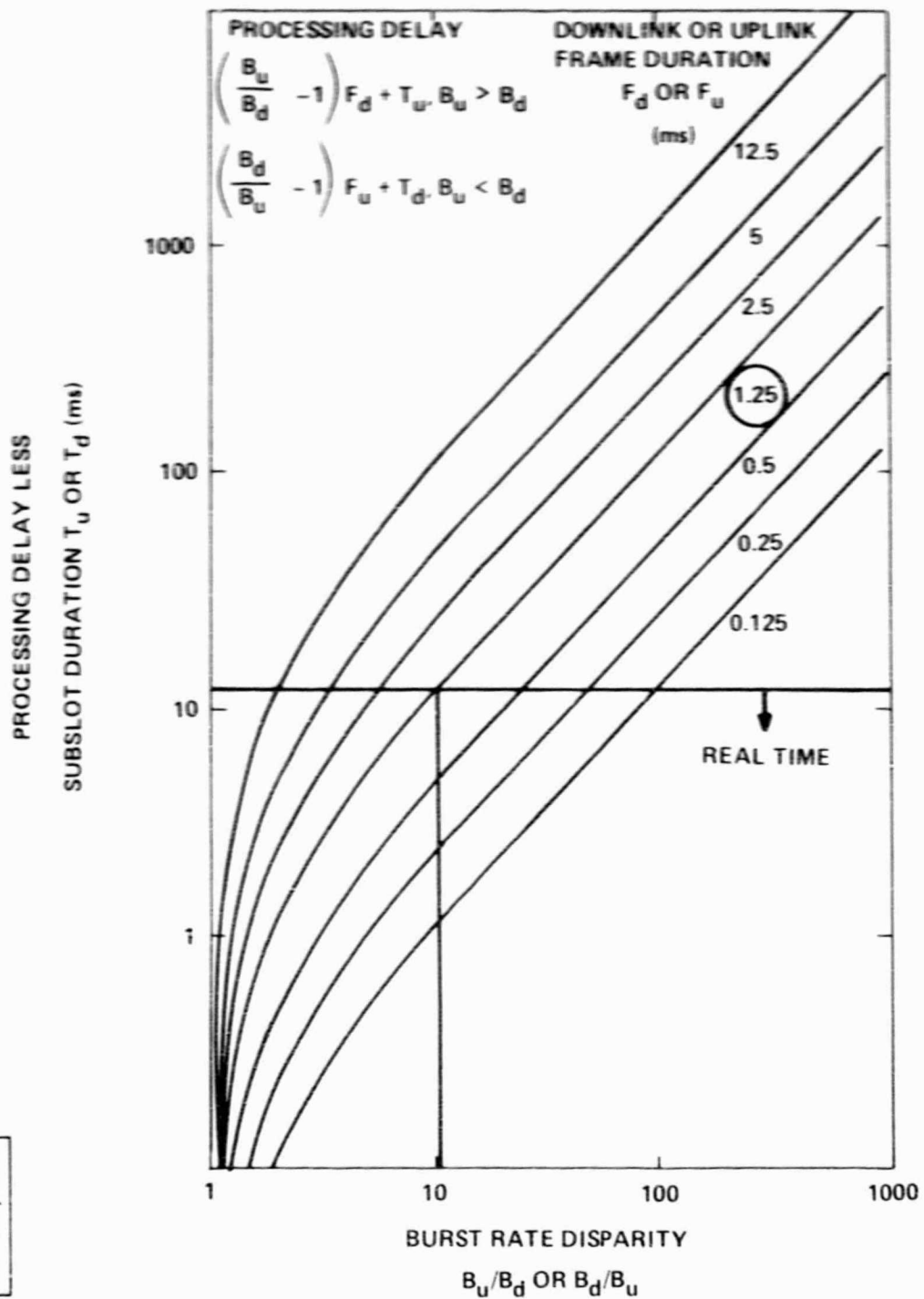


Figure 3-17. On-Board Processing Delay for Disparate Uplink and Downlink Burst Rates

and

$$A_d = N_{fd} + M_{sd}F_{sd}/T_{sd} \quad (3.27b)$$

These numbers could be in the order of 100, at least, to provide sufficiently discriminatory coverage patterns for the United States. At 30/20 GHz, 400 slightly overlapping beam areas would cover the U.S. with one beam area between the Boston and New York beams and two beam areas between the San Francisco and Los Angeles beams [Staelin, 1976].

For digital voice or real time data, the maximum frame interval must be comparable to a round trip propagation delay. Otherwise, too much delay would be introduced for commercial acceptance. Let  $F$  be no more than 1/4 sec. Without any queueing delays or processing delays due to burst rate disparities, this implies a total turn-around delay of at most 5/4 sec, including two frames' worth of demodulations at the satellite and one frame's worth at each of the two terminals involved in the conversation.

It is unlikely that the total frequency allocation would exceed a few GHz at extremely high frequency (EHF) (30 GHz and above) [Jain, 1979].\* In terms of interbeam isolation, it should not be difficult to guarantee complete frequency reuse for each scanning beam, since it would be unnecessary to ever have scanning beam patterns that were overlapping or adjacent simultaneously. It may also be possible to satisfactorily realize a fixed frequency plan

\* Current status:

30/20 GHz

- |                                   |                                 |  |
|-----------------------------------|---------------------------------|--|
| o fixed-satellite and terrestrial | 27.5 - 29.5/17.7-19.7 (2.0 GHz) |  |
| o fixed-satellite                 | 29.5-30/19.7-20.2 (0.5 GHz)     | WARC-79 recommendation for mobile-satellite, also. |
|                                   | 30-31/20.2-21.2 (1 GHz)         |  |
|                                   | government only                 |  |

41-43 GHz

- o broadcasting satellite (2 GHz)

43-48 GHz

- |  |                   |   |
|--|-------------------|---|
| o aeronautical and maritime mobile-satellite | DoD 43-45 (2 GHz) | WARC-79 recommendation for mobile-satellite in 43-45 GHz, also. |
|  | DoT 45-48 (3 GHz) |   |

50/40 GHz

- |                   |                     |   |
|-------------------|---------------------|---|
| o fixed-satellite | 50-51/40-41 (1 GHz) | WARC-79 recommendation for terrestrial service, also. |
|-------------------|---------------------|---|

with a frequency reuse factor  $M_f$  with no beam overlaps, or  $M_f/3$ , with beam overlaps. In order to better accommodate a close packing of fixed beams in densely populated areas, it may be necessary to adopt the latter three frequency plan.

Bandwidth efficient modulations exist such that it is possible to space adjacent carriers by as little as  $3/4$  to  $3/2$  the larger burst rate apart even when the individual signals are mutually uncoordinated in symbol timing, frequency uncertainty, and transmitter powers. Thus, it seems safe to assume a bits/cycle packing factor  $\beta$  between  $2/3$  and  $4/3$ , at least. (See section 3.3)

If complete frequency reuse and disjoint frequency bands for fixed and scanning beams are assumed, so that a scanning beam can overlap a fixed beam without interference, the total instantaneous bandwidth becomes

$$\begin{aligned} W &= W_f + W_s \\ &= B_f N_f / \beta_f + B_s N_s / \beta_s \end{aligned} \quad (3.28)$$

where  $B$  and  $N$  denote burst rate and number of carriers in a beam. This overall bandwidth  $W$  cannot exceed the allocation, e.g., 2.5 GHz at 30/20 GHz.

Another approach that may be adopted is to assume the existence of frequency allocations in contiguous 500 MHz or 600 MHz segments, for example. That is, there could be anywhere from one to five 500 MHz sub-bands or from one to four 600 MHz sub-bands allocated to a given beam, depending on frequency management issues. A sub-band larger than 600 MHz for a single carrier may be difficult to justify with state-of-the-art technology related to maximum achievable burst rates, TWT dispersion effects, etc., or dispersions in time and frequency induced by the propagation medium. Thus, a basic 500 MHz or 600 MHz allocation bandwidth unit permits a flexible, modular, technologically feasible design.

Digital Switch Implementation. The  $m \times m$  digital circuit switch size is determined by number of simultaneous T1 and T2 user pairs, i.e.,

$$m = m_1 + m_2 = N_1 N_1 + N_2 N_2. \quad (3.29)$$

A switch of order  $m = 36$  is implied by  $M_1 = 2$  scanning uplink beams with  $N_1 = 8$  carriers per beam for the T1 users and  $M_2 = 4$  scanning uplink beams with  $N_2 = 5$  carriers per beam for the T2 users, for

example. The downlink mixture could be different, e.g.,  $M_1 = 1$ ,  $N_1 = 16$ ,  $M_2 = 2$ ,  $N_2 = 10$ , provided the number of downlink and uplink users remain equal for real time point-to-point communications. Such a switch should be easily realizable with today's large scale integration (LSI) circuitry.

For example, a digital feed forward network of 36 OR gates of 36 inputs each (or 252 OR gates of 6 inputs each) and  $36 \times 36 = 1296$  AND gates of 2 inputs each with one control line per AND gate is sufficient. Alternatively, a  $36 \times 36$  switch could be implemented by a non-blocking Clos network composed of six  $6 \times 11$  switches, six  $11 \times 6$  switches and eleven  $6 \times 6$  switches, (1188 2-input AND gates and 240 6-input (at most) OR gates).

If some blocking probability can be tolerated, a  $36 \times 36$  switch could be implemented as a two-stage network of twelve  $6 \times 6$  switches, where each  $6 \times 6$  switch is composed of six OR gates and 36 AND gates, again with one control line per AND gate. This totals  $12 \times 6 = 72$  OR gates of only six inputs each, and only  $12 \times 36 = 432$  AND gates and control lines. Still another realization of a  $36 \times 36$  switch would be a 3-stage network of twenty-four  $3 \times 3$  switches and nine  $4 \times 4$  switches totaling  $24 \times 3 + 9 \times 4 = 108$  OR gates and  $24 \times 3 \times 3 + 9 \times 4 \times 4 = 360$  AND gates and control lines. This realization would probably have a smaller blocking probability than the 2-stage network. Specific digital switching designs of this sort are outlined in appendix A.

One should note that integrated circuit chips with roughly half a million gates are becoming feasible with  $1/2$  micron electron beam lithography fabrication techniques [Gossen, 1979]. This suggests that brute force designs that are easy to test (due to a homogeneous layout, for example) will become even more practical. Designing to minimize the number of gates per function is becoming less and less important. Conceivably, one might opt for a single stage switching design even if the number of input/output pairs approaches 1000.

#### 3.2.2.4 Design Examples

An initial attempt at a theoretical optimization of a multiple beam satellite overlay model for a real time point-to-point terrestrial network is contained in appendix B. The examples of this subsection are based on that analysis. Other approaches are also of interest [McGarty, Warner, 1977].

In this model the satellite users are partitioned into four classes according to the data rates of the TX carrier system. The

T1 and T2 users are assumed to operate in an FDMA/burst mode in scanning beams while the T3 and T4 users operate in a TDMA mode in fixed beams. More than one TDMA carrier per beam is allowed in a fixed beam. Only the T1 and T2 user signals are regenerated in the satellite; the T3 and T4 signals are RF switched. The signal bandwidths occupied by the four user classes are assumed to be disjoint and contained within a 2.5 GHz allocation.

The results of appendix B are summarized as follows. For a fixed number of simultaneous beams  $M$ , a fixed number of coverage areas  $A$ , and a fixed total signal bandwidth  $W$ , a critical point for the total satellite throughput  $t$  occurs for

$$M_i = M/\beta_i \sum_{j=1}^4 \beta_j^{-1}, \quad i = 1, 2, 3, 4 \quad (3.30)$$

$$W_2/W_1 = \beta_1/\beta_2 \quad (\text{assuming } r_2 = r_1) \quad (3.31a)$$

$$W_4/W_3 = \beta_3/\beta_4 \quad (3.31b)$$

where  $M_i$ ,  $\beta_i$ ,  $W_i$  and  $r_i$  are the number of beams, bps per Hz, bandwidth, and burst to data rate ratio for the users of group  $i$ , respectively. The complexity of a cross-multiplexing digital switch of size  $\underline{m} = m_1 + m_2$  is critical if

$$M_2 = M_1 \quad (3.32a)$$

$$N_2 = N_1 \quad (3.32b)$$

where  $N_i$  is the number of simultaneous FDMA carriers per scanning beam. Without cross-multiplexing the complexities of the digital switches are minimized if

$$N_2/N_1 = (M_1/M_2)^{1/2}. \quad (3.33a)$$

Similarly, the complexities of the RF switches are minimized if

$$N_4/N_3 = (M_3/M_4)^{1/2}. \quad (3.33b)$$

It also appears desirable to select group burst lengths according to

$$T_2 = T_1 \quad (3.34a)$$

$$T_4 = T_3. \quad (3.34b)$$

Equations (3.30) to (3.34) should be used only as guidelines to reasonable designs, not as representing a global optimization, because the system model does not yet include power budgets and more

direct indicators of satellite and terminal costs. Nevertheless, these results form a theoretical basis for subsequent optimization.

The relationship among fundamental system parameters is shown by the design example of figure 3-18. Suppose the system constraints include  $A = 100$  coverage areas,  $M = 24$  simultaneous beams, and  $W = 2.5$  GHz bandwidth. Assume that the burst to data rate ratios for the T1 and T2 users are equal, i.e., let  $r_2 = r_1$ . In addition, suppose the bits per cycle parameters are taken as  $\beta_1 = \beta_2 = 1/2$  and  $\beta_3 = \beta_4 = 1$ , indicating more efficient usage of the spectrum in the fixed beams. Finally, assume burst durations of  $T_1 = T_2 = T_3 = T_4 = 125 \mu\text{sec}$ . Then according to the flow of figure 3-18, it is seen that the following parameters result:

$$M_1 = M_2 = 8 \quad M_3 = M_4 = 4 \quad (3.35a)$$

$$N_1 = 3 \quad N_2 = 1 \quad N_3 = N_4 = 2 \quad (3.35b)$$

$$r_1 = r_2 = 6 \quad r_3 = 13 \quad r_4 = 2 \quad (3.35c)$$

$$A = 104 \quad (3.35d)$$

$$B_1 = 9.26 \text{ Mb/s} \quad B_2 = 37.9 \text{ Mb/s}$$

$$B_3 = 581 \text{ Mb/s} \quad B_4 = 552 \text{ Mb/s} \quad (3.35e)$$

$$W_1 = 55.6 \text{ MHz} \quad W_2 = 75.7 \text{ MHz} \quad W_3 = 1.16 \text{ GHz}$$

$$W_4 = 1.10 \text{ GHz} \quad W = 2.4 \text{ GHz} \quad (3.35f)$$

$$F_1 = F_2 = 750 \mu\text{sec} \quad F_3 = 1.625 \text{ msec} \quad F_4 = 250 \mu\text{sec} \quad (3.35g)$$

The equations of (3.35a) resulted from invoking equation (3.30), but this is not necessarily required in a design procedure. Although equations (3.31a) and (3.31b) were used, an independent decision is necessary to determine the relative direct-to-user (T1 and T2) and trunking (T3 and T4) bandwidths, initially. In this example, only 100 MHz is nominally allocated to the T1 and T2 users in deference to possible limitations of on-board memory speeds. For convenience of implementation, the  $r_i$ 's must be integers, and rounding is used. The  $r_1 = r_2$  value is determined by a generalization of equation (3.27) given in appendix B. Note that this perturbs  $A$  to 104. The burst rates and numbers of carriers per beam for the T1 and T2 users are determined by equation (3.21) and an analog of equation (3.28) found in appendix B. Since  $N_1$  and  $N_2$  result from rounding to the nearest integers, the signal bandwidths



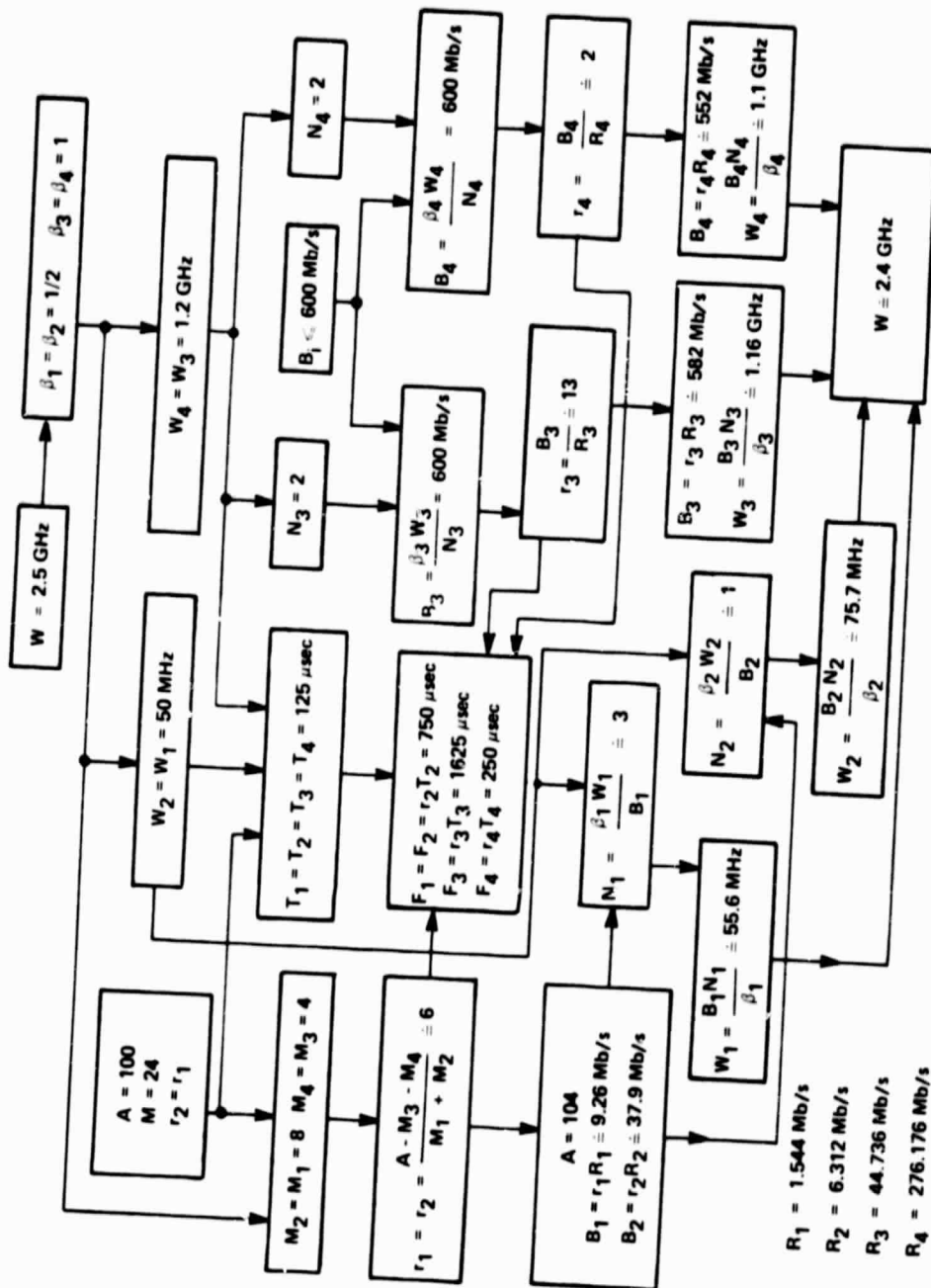


Figure 3-18. Example Design Flow and Parameter Relationships

$W_1$  and  $W_2$  are perturbed from their initial values. The burst rates  $B_3 = B_4 = 600$  Mb/s resulted from jointly selecting  $\beta_3 = \beta_4 = 1$ ,  $W_3 = W_4 = 1.2$  GHz, and  $N_3 = N_4 = 2$ , subject to the maximum burst rate constraint mentioned previously. The integer values of  $r_3$  and  $r_4$  follow from equation (3.21), which perturbs  $B_3$ ,  $b_4$ ,  $W_3$ , and  $W_4$ . Thus, a total signal bandwidth of  $W = 2.4$  GHz results which is slightly less than the original 2.5 GHz allocation.

Derived parameters for this design example are shown in figure 3-19. For these  $M_i$ 's, note from equations (3.33a) and (3.33b) that the complexities of the digital switches are not minimized because  $N_2 \neq N_1$ , while those for the RF switches are minimized since  $N_4 = N_3$ . The throughputs and numbers of processing bits for the T1 and T2 users are quite reasonable for on-board regeneration. A healthy total throughput of 9.6 Gb/s is attained for a total of 312 users. The selected burst duration of 125  $\mu$ sec results in frame intervals amenable to real time data communications even with some uplink and downlink burst rate disparities.

In this first example, the number of scanning beams exceeded the number of fixed beams because equation (3.30) was used to compensate for the inferior bandwidth efficiencies assumed for the scanning beams. It may be desirable to reduce the relative number of scanning beams to simplify the spacecraft. Toward this end, two more examples are considered under the assumption of equal bandwidth efficiencies, viz.,  $\beta_i = 1$ . Equations (3.30) to (3.34) were applied to obtain the results of figure 3-20. Figure 3-21 shows results with still fewer scanning beams after applying only equations (3.31) and (3.33b).

Group Parameter	T1 $i = 1$	T2 $i = 2$	T3 $i = 3$	T4 $i = 4$	Total
Simultaneous Users $m_i = M \cdot N_i$	24	8	8	8	48
Concurrent Users $n_i = m_i \cdot r_i$	144	48	104	16	312
Throughput $t_i = n_i \cdot R_i$	222 Mb/s	303 Mb/s	4.65 Gbps	4.42 Gb/s	9.6 Gb/s
Bits per User per Frame $b_i = B_i \cdot T_i$	1,158	4,734	72,696	69,044	
Processing Bits $q_i = m_i \cdot b_i$	27.8 K	37.9 K	582 K	552 K	1.2 M
Control Lines $m_i^2$	576	64	64	64	

$$\begin{array}{llll}
 A = 104 & r_3 = 13 & r_4 = 2 & W = 2.4 \text{ GHz} \\
 M = 24 & M_1 = M_2 = 8 & M_3 = M_4 = 4 & \beta_1 = \beta_2 = 1/2 \quad \beta_3 = \beta_4 = 1 \\
 r_1 = r_2 = 6 & N_1 = 3 & N_2 = 1 & N_3 = N_4 = 2 \\
 & & & B_i \leq 600 \text{ Mb/s} \\
 & & & T_i = 125 \mu\text{sec} \\
 & & & F_1 = F_2 = 750 \mu\text{sec} \\
 & & & F_3 = 1.625 \text{ msec} \\
 & & & F_4 = 250 \mu\text{sec}
 \end{array}$$

Figure 3-19. Derived Parameters for Design Example

Group Parameter	T1 $i = 1$	T2 $i = 2$	T3 $i = 3$	T4 $i = 4$	Total
Simultaneous Users $m_i = M \cdot N_i$	12	12	12	12	48
Concurrent Users $n_i = m_i \cdot r_i$	84	84	156	24	348
Throughput $t_i = n_i \cdot R_i$	130 Mb/s	530 Mb/s	7.0 Gb/s	6.6 Gb/s	14.3 Gb/s
Bits per User per Frame $b_i = B_i \cdot T_i$	1,351	5,523	72,696	65,444	
Processing Bits $q_i = m_i \cdot b_i$	16,212	6,276	872 K	829 K	1.78 M
Control Lines $m_i^2$	144	144	144	144	

$$\begin{aligned}
 A &= 96 & W &= 2.4 \text{ GHz} & T_i &= 125 \mu\text{sec} & F_1 &= F_2 = 875 \mu\text{sec} \\
 M &= 24 & M_i &= 6 & N_i &= 2 & F_3 &= 1.625 \mu\text{c} \\
 r_1 &= r_2 = 7 & \beta_i &= 1 & B_i &\leq 600 \text{ Mb/s} & F_4 &= 250 \mu\text{sec}
 \end{aligned}$$

Figure 3-20. Results of Sample Optimization

Group Parameter	T1 $i = 1$	T2 $i = 2$	T3 $i = 3$	T4 $i = 4$	Total
Simultaneous Users $m_i = M_i N_i$	8	2	20	20	50
Concurrent Users $n_i = m_i r_i$	160	40	260	40	500
Throughput $t_i = n_i R_i$	247 Mb/s	252 Mb/s	11.6 Gb/s	11.0 Gb/s	23.2 Gb/s
Bits per User per Frame $b_i = B_i T_i$	3,860	15,780	72,696	69,044	
Processing Bits $q_i = m_i b_i$	30,880	31,560	1.45 M	1.38 M	2.9 M
Control Lines $m_i^2$	64	4	400	400	

$A = 100$        $W = 2.5 \text{ GHz}$        $T_i = 125 \mu\text{sec}$        $F_1 = F_2 = 875 \mu\text{sec}$   
 $M = 24$        $M_1 = M_2 = 2$        $M_3 = M_4 = 10$        $F_3 = 1.625 \text{ msec}$   
 $r_1 = r_2 = 20$        $\beta_i = 1$        $N_1 = 4$        $N_2 = 1$        $N_3 = N_4 = 2$        $F_4 = 250 \mu\text{sec}$        $B_i \leq 600 \text{ Mb/s}$

Figure 3-21. Results with Fewer Scanning Beams

### 3.3 MODULATION/DEMODULATION

For direct-to-user and trunking services at TX carrier data rates, modulation issues are less critical than for small user terminals. The main theme of this section is bandwidth efficiency.

With TDMA and sufficient guard time between user bursts, interchannel interference of the modulation scheme is irrelevant because the entire satellite transponder is dedicated to only one signal at a time. However, even with TDMA and a nonlinear repeater, the choice of modulation can make a difference in throughput data rates and intersymbol interference (ISI) effects. The type of digital modulation can also be an important factor in the performance of signal acquisition and tracking procedures for on-board demodulation, especially for short TDMA bursts where the signal overhead induced by preamble can be significant. The relatively long bursts expected for the TX carrier systems make acquisition and tracking less of an issue.

With FDMA, uncoordinated user transmissions, and on-board regeneration, improved digital modulations can have a dramatic impact on the number of uplink channels that can be effectively packed into a given section of bandwidth. Results developed in the last few years show how to achieve an order of magnitude or more improvement in crosstalk levels by slight modifications of waveform structure that are quite reasonable to implement with today's integrated circuit (IC) technology.

Bandwidth efficient digital modulation is becoming especially important for the on-board regeneration of signals from disadvantaged user terminals. With FDMA, the more appropriate access method for simple terminals, multifarious signals must be received simultaneously and with adequate reliability at the satellite. The problem is compounded by the possibilities of different relative received power levels, burst rates, symbol timing, and frequency shifts from nominal preassigned values. The network control problem of precisely coordinating the transmissions of all terminals could be formidable and prohibitively expensive, except for those terminals which are clearly not disadvantaged and that can operate best in a TDMA circuit switched mode. Although the T1 and T2 direct-to-user terminals tend to belong to the latter class, low crosstalk modulations are emphasized in this section for their potential application to T1 and T2 users organized in an FDMA packet switched fashion.

Bandwidth efficiency is currently receiving more attention even in analog signaling domains, involving frequency modulation (FM) in particular [Lusignan, 1978]. It is conceivable that analog voice modulations will become very attractive candidates for uplinks to regenerating satellites acting as intermediate interoperability nodes for improving connectivities among lower cost users. However, this section concentrates on digital modulations because of the emphasis on highly reliable data communications applications in this report.

Digital, constant envelope modulations, which are especially important because of their efficiency for use with peak power limited transmitters, are highlighted. Pseudonoise (PN), chirp, frequency hopping (FH), and hybrid spectral spreading variations are not covered. Frequency management issues associated with particular user services should be considered further in making recommendations for technology development.

Consider digital modulations appropriate for satellite terminals bursting in an FDMA mode while in the footprint of a scanning uplink beam. For these relatively small, remotely located terminals, the objective is to accomplish bandwidth efficiency even though no network timing, frequency, or power control discipline may be imposed. When such uncoordinated FDMA signals arrive at the satellite, significant undesirable crosstalk may result unless the non-orthogonality of the signals is compensated by some sort of filtering.

Filtering may be accomplished at a terminal, at the satellite, or both. Adequate filtering of conventional modulations at the transmitter may be difficult. The use of pre-shaped modulation waveforms with attractive spectral roll-off properties may be advantageous. However, baseband shaping at the receiver can further reduce crosstalk; the transmission spectrum does not tell the whole story, especially if the receiver is intentionally mismatched to the transmitter [Kalet, White, 1977]. This conclusion is illustrated by figures 3-22 and 3-23 which can be used to compare the relative spectral efficiency of several modulations from bandwidth and crosstalk points of view, respectively.

The higher the degree of continuity in the baseband waveforms, the faster the asymptotic roll-off of the crosstalk, i.e., the unwanted output of an integrate-and-dump demodulator [Reiffen, White, 1978]. Unfortunately, there is no unifying theory, which predicts relative crosstalk levels at the closer center frequency separations [Eaves, Wheatley, 1979]. However, computer simulations have shown that one form of continuous phase FSK known as MSK and

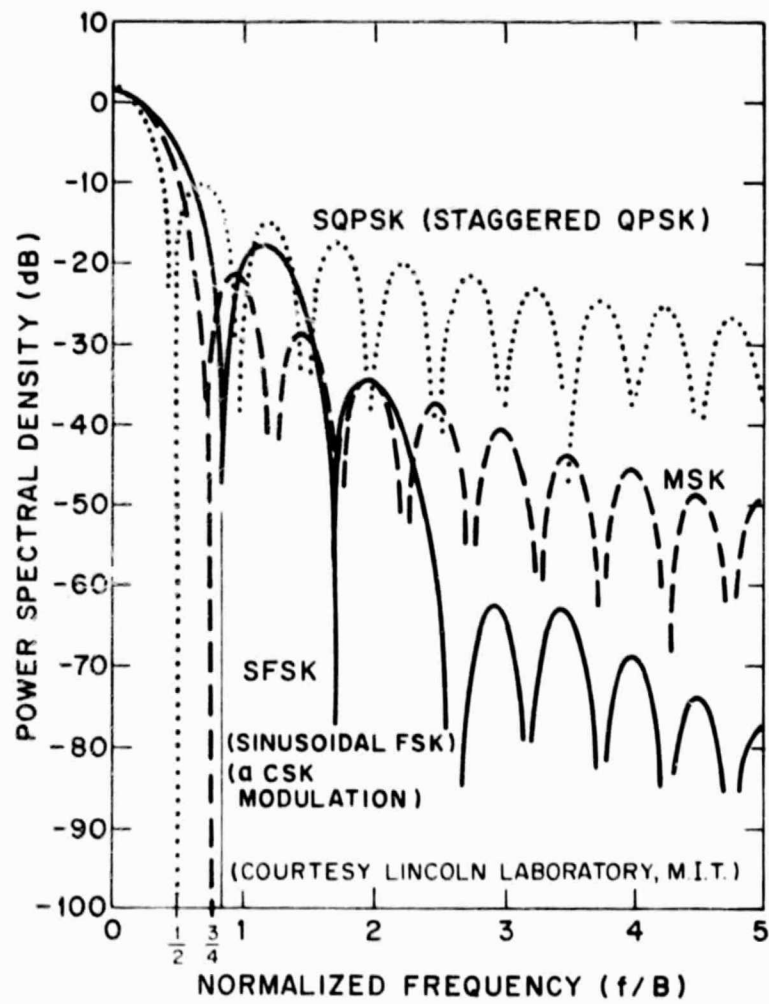


Figure 3-22. Spectral Densities of SQPSK, MSK, and SFSK



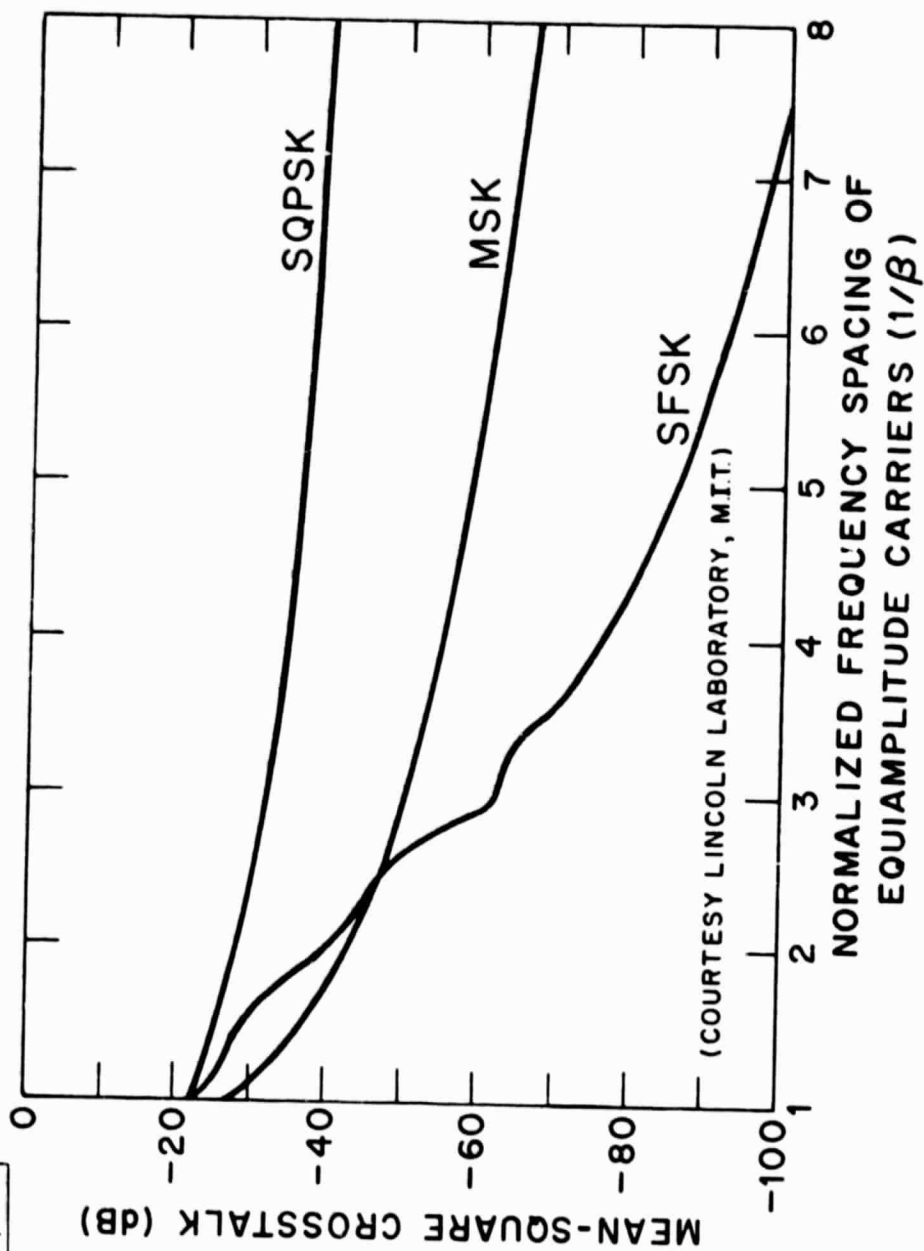


Figure 3-23. Asymptotic Behavior of Mean-Square Crosstalk

some of its generalizations to continuous (frequency) shift keying (CSK) have excellent bandwidth efficiency compared with QPSK modulations, for example [White, 1977]. With unsynchronized signals which vary in received power by as much as 10 or 15 dB, a weak signal can still be reliably demodulated with roughly five to ten times as many CSK signals per unit bandwidth compared to QPSK.

Stated another way, a typical center frequency separation for CSK signals is  $3/2$  to 2 times the burst rate. This can be improved upon by forming a weighted sum of adjacent data bits to produce very smooth phase transitions [de Jager, Dekker, 1978]. With this so-called tamed FM signaling, separations of only about  $3/4$  the burst rate may be feasible. Tamed FM requires only about a 1 dB larger  $E_b/N_0$  for a given bit error probability in additive white Gaussian noise (AWGN) than CSK which performs with the coherent detection\* efficiency of antipodal signals. In bandlimited situations, however, such a small increase in the energy contrast ratio required may be well worth the increased packing density attainable.

The salient parameter is the number of bits/cycle or bps per hertz, which for FDMA is defined as the channel burst rate  $B$  normalized by the center frequency separation  $\Delta f$ , i.e.,  $\beta = B/\Delta f$ . Although used freely in the multiple access/multiplexing section of the report, this parameter is meaningful only when associated with particular performance levels and specified constraints, or lack thereof, on network symbol timing or transmitter power control. A complete comparison among all modulation schemes mentioned is currently impossible because of gaps in available data. However, it is possible to state with assurance that low crosstalk digital modulations exist with  $\beta > 2/3$ , e.g., phase comparison sinusoidal frequency shift keying (PCSFSK) and tamed FM, which perform within 1 dB of the ideal even when the received FDMA signals are mutually unsynchronized and vary in power by up to 10 or 15 dB. Equivalent performance with  $\beta$  as large as 2 may eventually be shown to be feasible. For now it may be conservatively assumed that  $\beta = 2/3$  for the digital modulation employed for FDMA.

---

\* CSK can also be detected non-coherently without much degradation in efficiency or crosstalk performance [White, et al., 1977; Metzger, 1978].

The close packing property is important not only to conserve the spectral resource but also to reduce digital sampling rates required to perform the demodulation of groups of FDMA signals. As an illustrative example, suppose the maximum feasible clocking rate is 300 MHz. (This is consistent with the discussion in the multiple access/multiplexing section related to a maximum QPSK burst rate of 600 Mbps.) For digital quadrature carrier modulations like QPSK and CSK, this implies a maximum complex sampling rate of 300 MHz at a satellite demodulator, i.e., the bandwidth of the composite received signal to be demodulated must not exceed  $W = 300$  MHz. For a packing factor of  $\beta = 2/3$ , the maximum product of burst rate  $B$  and number of FDMA carriers  $N$  is  $BN = \beta W = 200$  Mb/s. Suppose the burst to data ratio for the T1 and T2 users is  $r = B/R = 20$ . Then, at most  $N = 200 \text{ Mb/s} / 20R$ , or about six T1 carriers or 2 T2 carriers per beam are possible. (A set of design curves is provided in figure 3-24. The required number of demodulators is obtained by multiplying by the number of beams for each user class.) If the packing factor  $\beta$  were smaller by a factor of 5, as it might be for unsynchronized QPSK, FDMA operation would be infeasible for an equivalent level of crosstalk performance. On the other hand, if  $\beta$  could be increased to 2, 19 T1 carriers and 5 T2 carriers per beam can be processed.

Some excellent work in continuous phase FSK (CPFSK) for obtaining coding-like performance without the introduction of redundant symbols and the attendant bandwidth increases is proceeding in Sweden [Rydbeck, Sundberg, 1979]. Researchers there are claiming effective transmission bandwidths which are just half those of tamed FM, although this does not necessarily imply tolerable center frequency separations of only  $3/8$  the burst rate. This may be feasible by making use of the intersymbol signaling constraints imposed on the baseband modulation for controlling spectral occupancy and crosstalk. If one plots a phase trellis for these modulations, it is apparent that a form of decoding compatible with the Viterbi algorithm is possible [Forney, 1973].

There are two basic thrusts to the Swedish research. The first is in the area of M-ary CPFSK where the baseband shaping is applied to only one data symbol [Aulin, Sundberg, 1979]. In the second approach, a binary CPFSK partial response system is considered where the shaping is applied over more than one signaling interval [Aulin, et al., 1979].

The characteristics and performance parameters of these and other modulations are covered in appendix C. A plot of bits/cycle versus detection efficiency for several modulations is shown there.

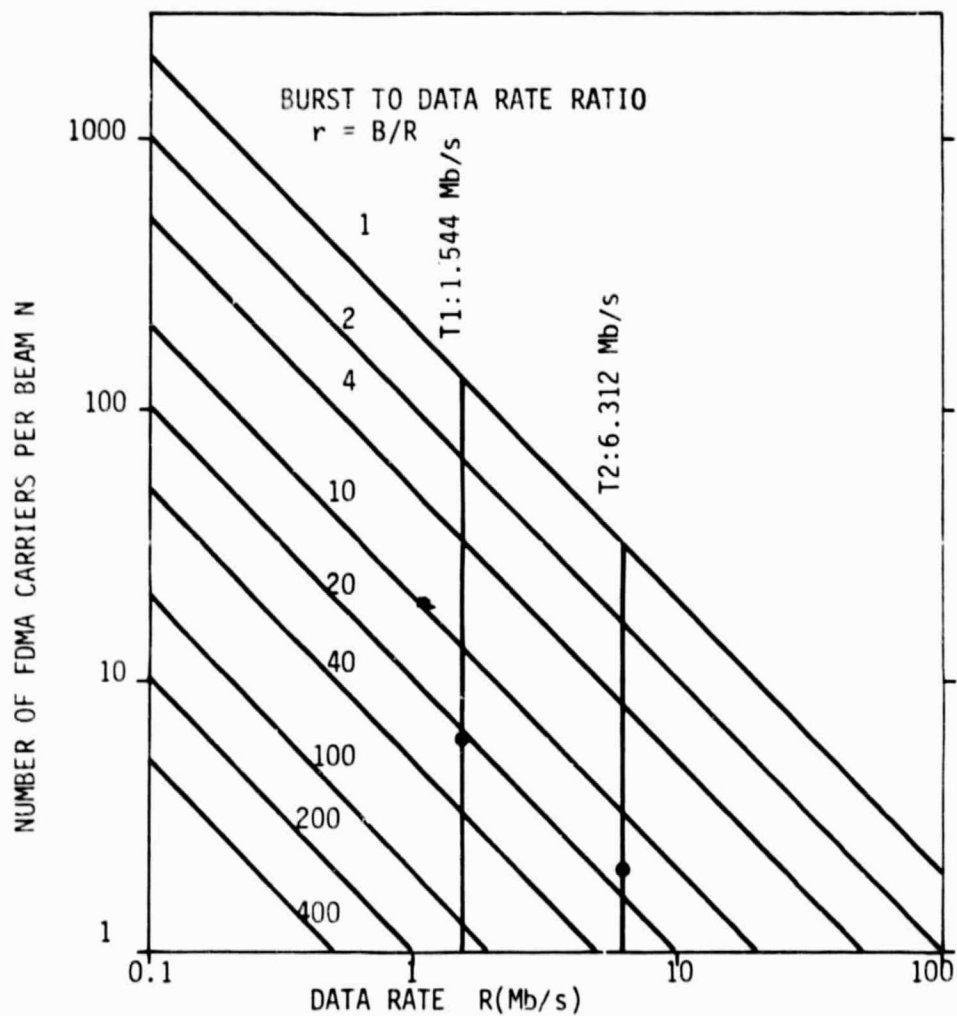


Figure 3-24. Number of Regenerators per Beam for Signal Bandwidth  $W = 300 \text{ MHz}$  and Carrier Packing Factor  $\beta = 2/3$  ( $\beta W = 200 \text{ Mb/s}$ )

### 3.4 REGENERATION

Modern digital terrestrial radio repeaters use the technique of demodulation, regeneration, and remodulation to improve the performance of terrestrial links. These regenerative repeaters operate at data rates from 40 Mb/s in Canada to 120 Mb/s in Great Britain, to 274 Mb/s in the U.S., to 405 Mb/s in Japan and provide reliable service for continuous QPSK transmission. For burst transmission of QPSK in a TDMA or scanning beam satellite system, the problems of demodulation and remodulation are complicated by the need to resynchronize the carrier phase and data clock of the demodulator during each burst period. Phase locked loop carrier reconstruction loops designed for continuous wave QPSK transmission do not work for burst transmission due to residual phase error. This has resulted in a new area of technology which has been specifically developed for burst carrier reconstruction for TDMA and scanning beam systems. Research in this area is well established; carrier synchronization techniques are being developed for burst data rates in the 1 Gb/s range. Although the impact of carrier reconstruction is included in our analysis of system performance, the actual implementation of carrier reconstruction techniques is not discussed.

#### 3.4.1 Repeater Concept

Regenerative repeaters are ideally suited for improving performance in digital satellite communications systems. The principal reason for the improvement is that the regenerative repeater, by virtue of signal detection, isolates the up and downlinks and thus avoids the direct transfer of noise accumulated on the uplink to the downlink. Rather, it passes on only errors made in the detection processing of the uplink signal and prevents signal anomalies occurring on the uplink from passing on to the downlink. The combination of detection and remodulation isolates, and hence permits, individual tailoring of the up and downlink modulation schemes such as FDMA uplink and TDM downlink. In addition, regeneration results in the elimination of the degradation caused by uplink filter induced AM, as well as the resulting PM of the QPSK carrier produced when the uplink signal is amplified directly by the nonlinear TWTA in the non-regenerating transponder design. Furthermore, remodulation permits the carrier to be derived from a common source on the satellite for all downlink signals and eliminates station to station and Doppler differences encountered as a result of repeating the uplink signal from different uplink stations. The resulting burst-to-burst carrier coherence at each receiver translates into simpler carrier reconstruction, improved

demodulator performance at the ground station, and less link degradation.

The end-to-end performance of the satellite communications link illustrated in figure 3-25 is used to quantify the various performance tradeoffs. This block diagram is a simplified version of a link through the satellite and is referred to as the system model. The link can handle either error correcting coded data or uncoded data with the link analysis. For example, in a regenerating link with no coding and using QPSK, each channel symbol contains two bits of information. In the demodulation process on board the satellite, the uplink channel symbols are restored to a noise free state in preparation for further processing and retransmission on the downlink. However, during demodulation some error may occur in the detection of uplink channel symbols which, for QPSK, results in two information bit errors. Since there is no correction mechanism, the resulting error is passed through the system to the downlink. We can see here then that the end-to-end link is simply a concatenation of two links with the effective end-to-end error probability a function of the signal-to-noise ratios in each link. This analysis is straight-forward; the comparative performance between a conventional repeater and a regenerative repeater for uncoded data is presented in the following section.

When data using error correcting codes are transmitted over a regenerating channel, the end-to-end analysis becomes more complex. This is because the coding process introduces into the data stream a specific type of redundancy which generally: (1) reduces the information data rate, and (2) spreads the bit information content over several channel symbols. It is the lack of a one-to-one relationship between information bits and channel symbols which complicates the analysis. For example, channel symbol errors may occur in the end-to-end link but no information bit error may occur as a result of the coding.

For coded systems in which the specific code is not specified, the link performance analysis can generally be done through an error bounding procedure. In the following sections, a performance measure based on error bounds is developed, and a conventional repeater and two types of regenerative repeater are compared.

In the analysis considered here, a conventional repeater is defined as a linear repeater with degradation added to account for non-linear elements in the system.

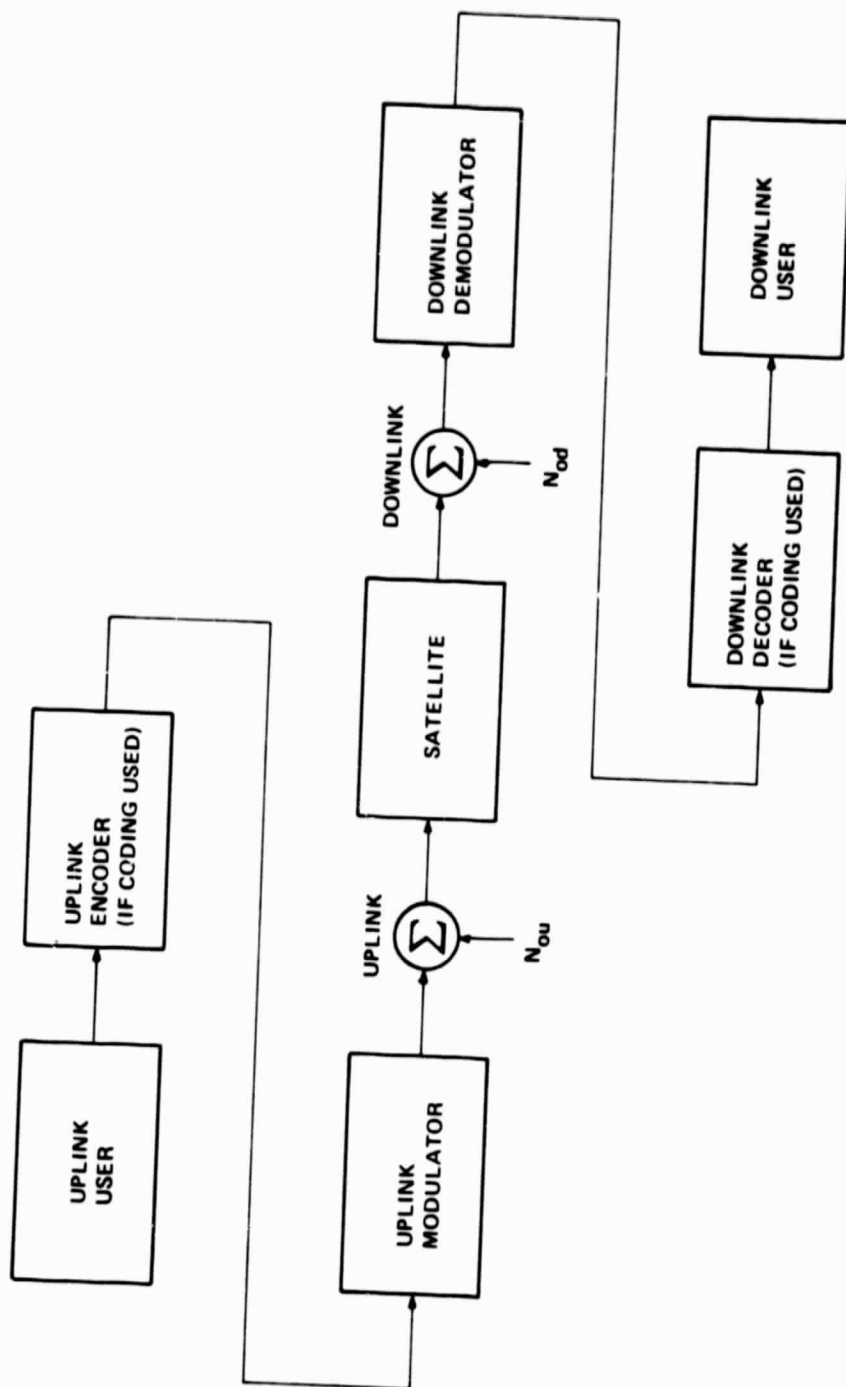


Figure 3-25. Satellite Communications System



### 3.4.2 Performance Analysis for Uncoded Data

Work by S. J. Campanella, COMSAT Laboratories, analyzed the performance of a digital regeneration technique for single-carrier per channel links which included the degradation in link performance due to ground station high power amplifier (HPA) and satellite traveling wave tube amplifier (TWTA) [Campanella, 1977]. The results for the uncoded case are presented here with the author's permission.

Measurements and experiments performed by COMSAT Laboratories have estimated the amount of degradation experienced in a satellite link for various backoff conditions resulting from the ground station high-power TWTA (HPA) and low power TWTA in the satellite. These results for a bit error rate (BER) of  $10^{-4}$  are given in table 3-1, indicating the degradation introduced by each device. These values, which will be used as a basis for the analysis which follows, indicate that the degradation encountered over the link increases significantly as the amount of backoff decreases in either the HPA of the earth station or the satellite TWTA. This degradation is caused in part by the interaction of AM (induced by the action of band limiting filters on the digitally modulated carrier) with the AM to PM nonlinearity component of the TWTA to produce PM of the RF carrier during the individual symbol periods. At the demodulator such PM degrades the performance of the coherent demodulation process and also interferes with carrier and bit timing recovery; this results in significant degradation relative to the back-to-back (perfect channel) modem performance.

Table 3-1

Link Degradation Due to the HPA and Satellite TWTA  
Nonlinearity for BER =  $10^{-4}$

HPA Output backoff (dB)	Satellite TWTA Backoff (dB)	Overall Degra- dation (dB)	HPA Degra- dation (dB)	Satellite TWTA Degrada- tion (dB)
6	4	2.8	1	1.8
6	0	3.2	1	2.2
0	0	5.2	3	2.2



To illustrate the advantage of on-board digital regeneration, uplink versus downlink  $E_b/N_0$  contours are calculated for constant bit error rate on a satellite link with and without on-board regeneration. To estimate the conventional (non-regenerative) performance, the up and downlink values of  $E_b/N_0$  are combined, using a linear assumption to obtain the effective  $E_b/N_0$  for the link as shown below. That is,

$$\left(\frac{E_b}{N_0}\right)_{\text{eff}} = \frac{E_{bu}/N_{ou} \quad E_{bd}/N_{od}}{E_{bu}/N_{ou} + E_{bd}/N_{od} + B/R} \quad (3.36)$$

where  $E_{bu}/N_{ou}$  and  $E_{bd}/N_{od}$  are the uplink and downlink signal energy per bit per noise power spectral density ratios respectively,  $B$  is the system bandwidth, and  $R$  is the data rate.

Calculation of a conventional repeater end-to-end link performance for coherent QPSK (CQPSK) uses the back-to-back modem performance represented by the theoretical curve in figure 3-26 plus an additional 1.8 dB modem implementation degradation and link degradation in table 3-1. The values of total degradation and the corresponding values of  $E_b/N_0$  needed to achieve a BER of  $10^{-4}$  for each combination of HPA and TWTA backoff examined are given in table 3-2. The resulting curves are shown in figures 3-27, 3-28, and 3-29 for 6/4, 6/0, and 0/0 HPA/TWTA backoff cases, respectively. These curves represent the performance presently expected with a conventional link and will serve as a basis for comparison with the performance of the same links operating via regeneration repeaters.

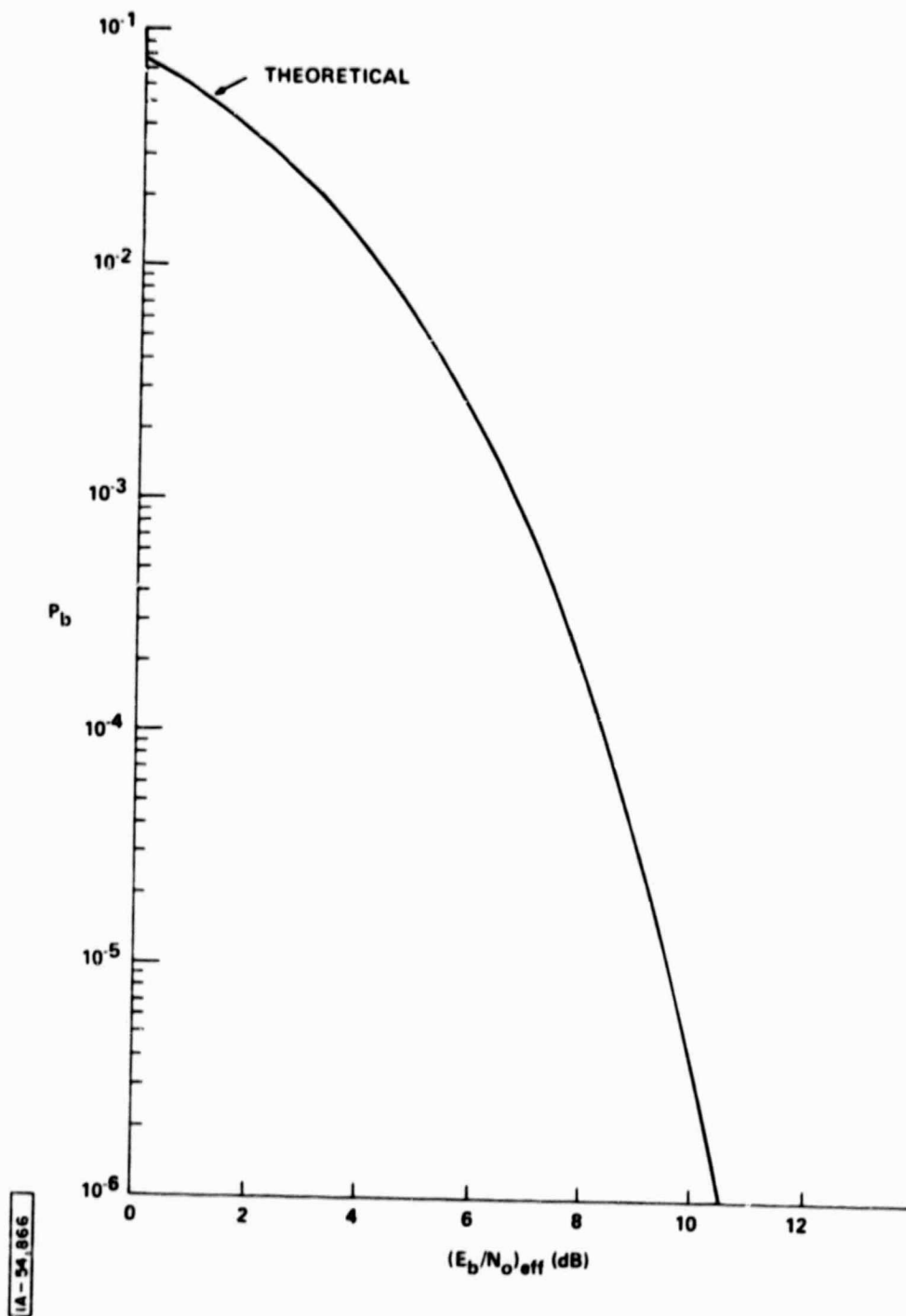


Figure 3-26. Probability of Error for QPSK

Table 3-2

Values of  $(E_b/N_o)_{\text{eff}}$  for Conventional Link Analysis  
at  $\text{BER} = 10^{-4}$

HPA Output Backoff (dB)	Satellite TWTA Backoff (dB)	Total Degrada- tion* (dB)	$(E_b/N_o)_{\text{eff}}$ (dB)
6	4	4.6	13.4
6	0	5.0	13.8
0	0	7.0	15.8

\* The values are dependent on both the group-delay and on amplitude characteristics of the filters in the channel.

To estimate the performance using a regenerative repeater in the link, it is necessary to determine the BER on the up and downlinks separately and to obtain the overall error rate by combining the up and downlink error rates as follows:

$$\text{BER}_U + \text{BER}_D = \text{BER}_{\text{eff}} \quad (3.37)$$

This expression ignores the influence of joint bit errors, which are of negligible consequence in the BER range of interest here.

For the uplink, differentially encoded CQPSK modulation is considered. The uplink degradation introduced by the HPA for each backoff case analyzed is given in table 3-1. An additional 1.8 dB modem implementation degradation must be included to obtain the total uplink degradation for each case. The resulting degradation values are given in table 3-3. Values of  $\text{BER}_U$  versus  $E_{bu}/N_{ou}$  are obtained by applying these degradation values to the performance curves of figure 3-26.

The degradation on the downlink is equal to the sum of the satellite TWTA degradation in table 3-1 less 1 dB to account for improvement due to the lessening impact of TWTA nonlinearities and 1.8 dB modem implementation margin. The resulting values are also given in table 3-3. Values of  $\text{BER}_D$  versus  $E_{bd}/N_{od}$  are obtained by applying the degradation values in table 3-3 to the performance curves of figure 3-26.

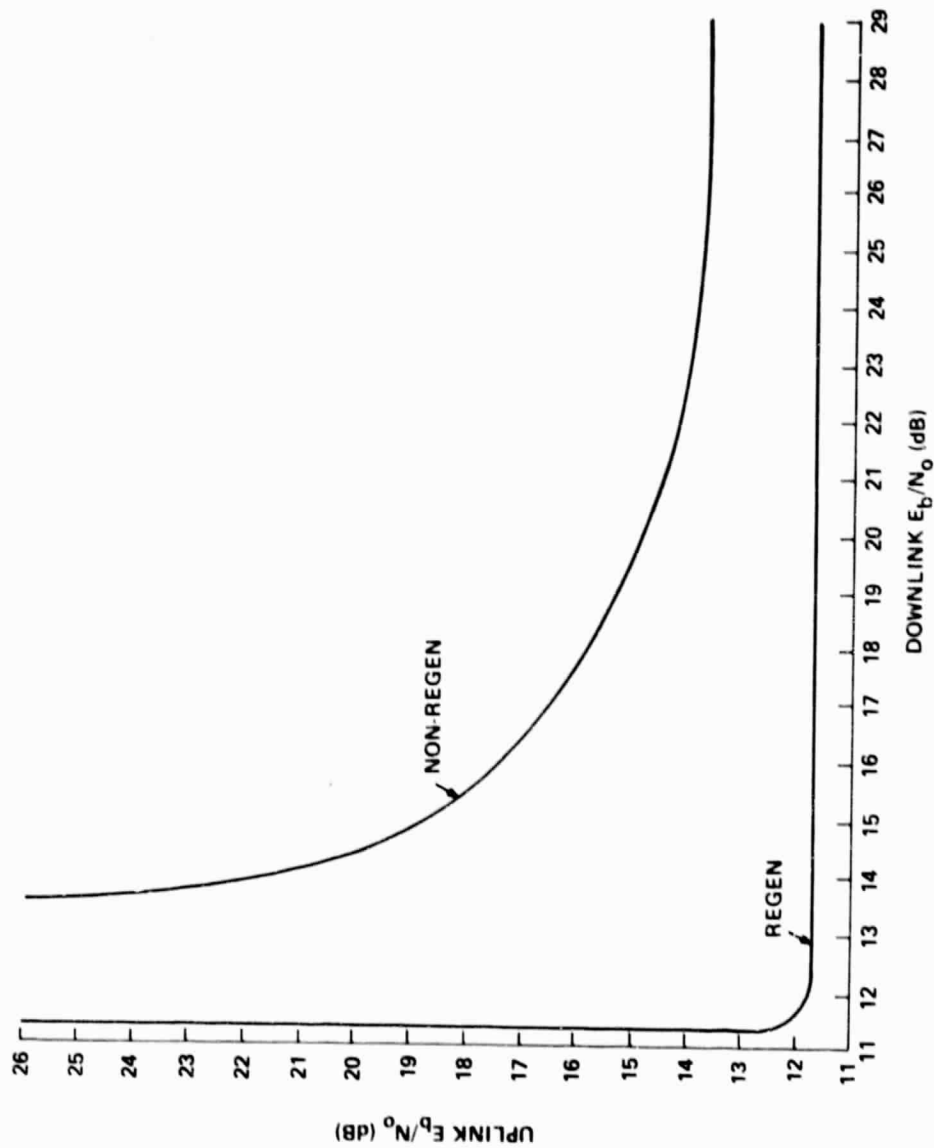


Figure 3-27. Performance Curves for Constant BER of  $10^{-4}$ ,  
HPA Backoff = 6 dB, TWT Backoff = 4 dB

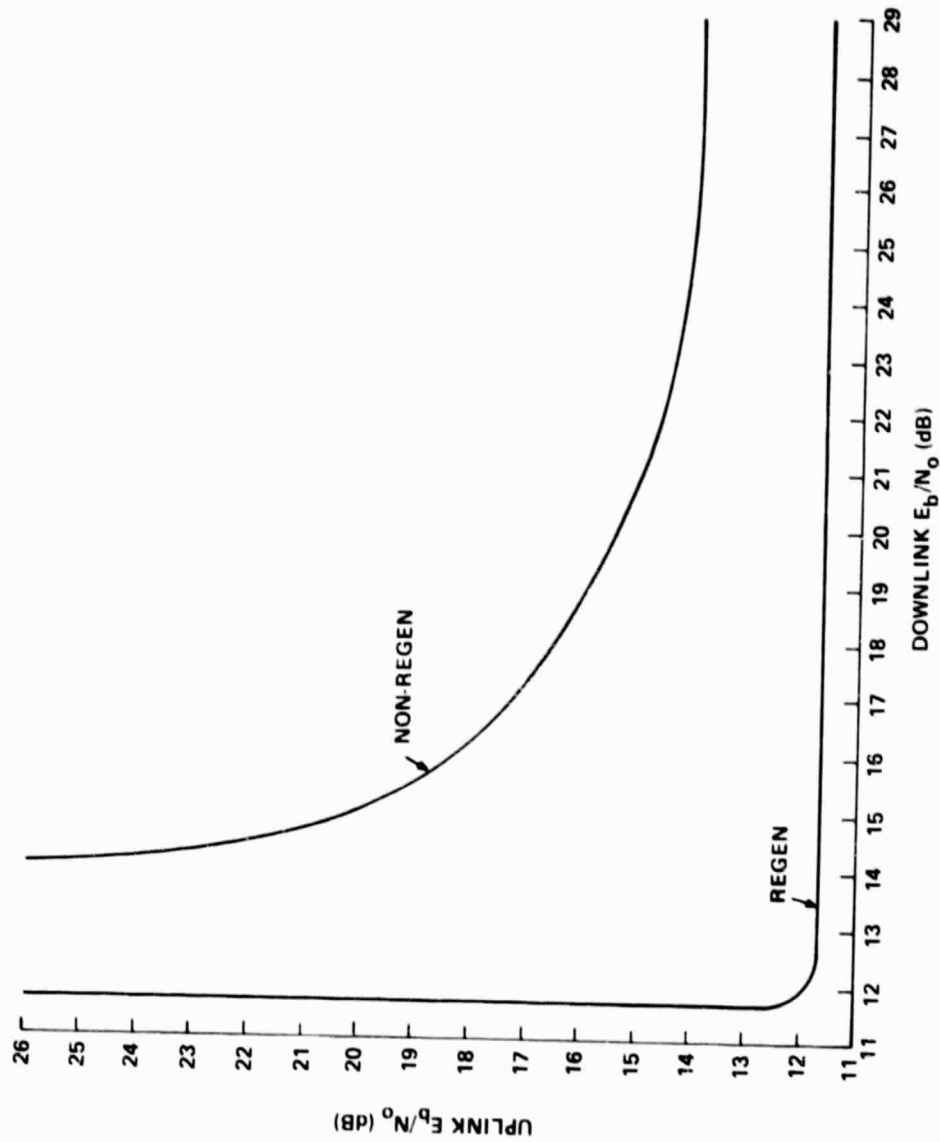


Figure 3-28. Performance Curves for Constant BER of  $10^{-4}$   
HPA Backoff = 6 dB, TWTA Backoff = 0 dB

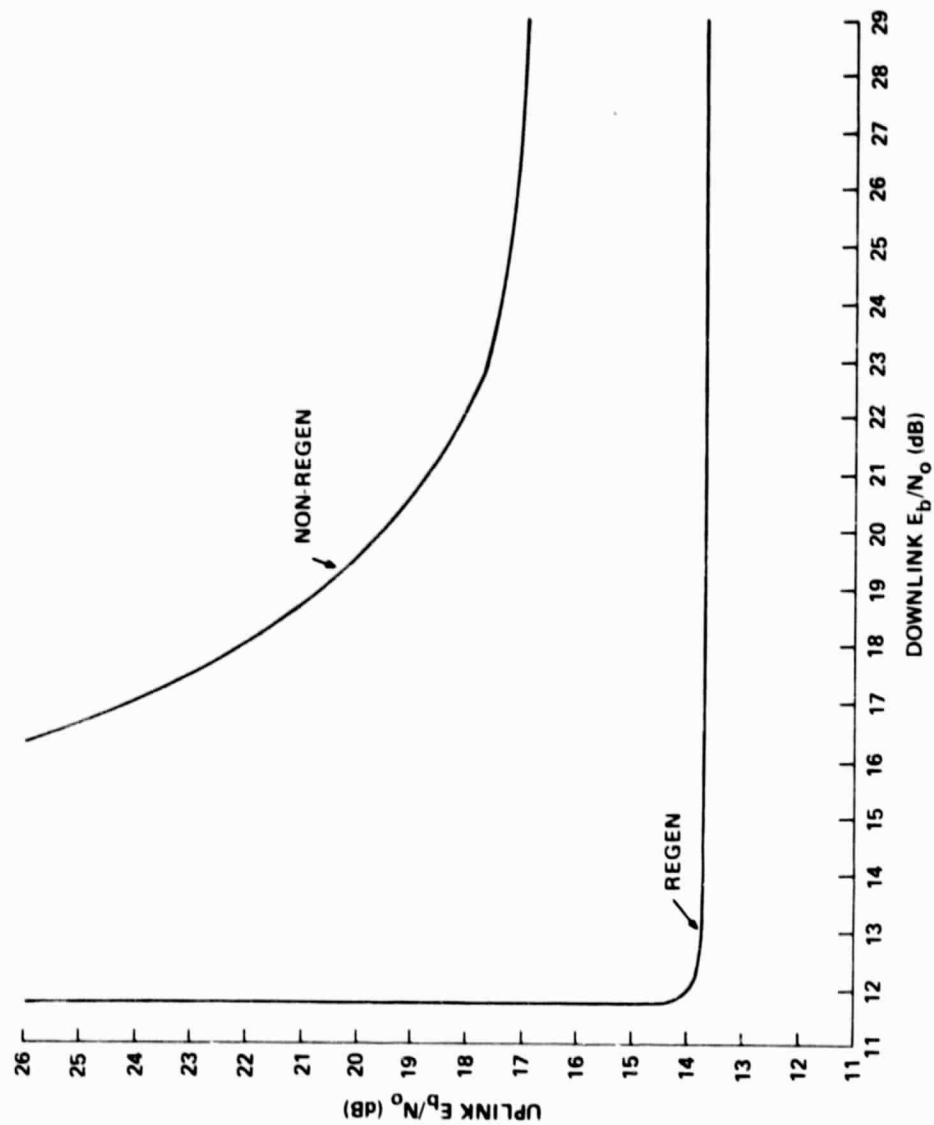


Figure 3-29. Performance Curves for Constant BER of  $10^{-4}$   
 HPA Backoff = 0 dB, TWTA Backoff = 0 dB

Table 3-3

Uplink and Downlink Degradation Used in the Regenerative  
Repeater Case

HPA Output Backoff (dB)	Satellite TWTA Backoff (dB)	Uplink Degra- dation (dB)	Downlink Degra- dation (dB)
6	4	2.8	2.6
6	0	2.8	3
0	0	4.8	3

The values are dependent on both the group delay and on amplitude characteristics of the filters in the channel.

The curves for the regenerative transponder have been obtained by using equation (3.37) to determine individual up and downlink BER values for an overall link bit rate of  $BER = 10^{-4}$ , using the BER versus  $E_b/N_0$  characteristics of figure 3-26, and appropriately applying the degradation of table 3-3. The results are plotted in figures 3-27, 3-28, 3-29, for 6/4, 6/0, and 0/0 HPA/TWTA backoff cases, respectively.

To determine the differences in the values of up and downlink  $E_b/N_0$  for the conventional and regenerative cases we will use points which are tangent to a line -1 slope line. Such points represent the joint minimization of the up and downlink values of  $E_b/N_0$  on the  $10^{-4}$  BER contours. The differences, which are a measure of the improvement due to regeneration, are listed in table 3-4 for the various backoff cases.

#### 3.4.3 Performance Analysis for Coded Data

In this analysis, the structures of the uplink encoder and downlink decoder are purposely left undefined. The downlink demodulator structure depends upon the type of processing done on-board the satellite. M-ary orthogonal modulation is assumed on the uplink. The modulator waveforms are each of single-sided bandwidth, B Hz, (centered around the carrier).

Table 3-4

Difference in Conventional and Regenerative Link  $E_b/N_0$   
Values for Uplink and Downlink

HPA Output Backoff (dB)	Satellite TWTA Backoff (dB)	Uplink Difference CQPSK (dB)	Downlink Difference CQPSK (dB)
0	0	5.0	6.8
6	0	4.8	4.8
6	4	4.6	4.6

In the system model, the uplink is assumed to contain signal and noise, the noise is spread uniformly over the system bandwidth  $B$ , and the downlink is assumed coherent and perturbed only by additive white Gaussian noise. On the uplink, noise is also generated both in the satellite receiver front end and from other sources. It has a one-sided spectral density height of  $N_{ou}$  watts/Hz. The uplink carrier-to-noise ratio at the satellite is  $C_u/N_{ou}$  and the received signal energy is  $E_u$ . On the downlink, noise is considered to be generated in the downlink receiver (apart from any retransmitted uplink noise). The downlink receiver noise has a single sided spectral density height of  $N_{od}$  watts/Hz. The downlink carrier-to-noise ratio is  $C_d/N_{od}$ ,  $C_d$  representing all power received at the ground from the satellite (both in the waveform and possibly in retransmitted uplink noise).

The on-board processing schemes considered for the coded case are in order of increasing complexity, (1) Hard Decision Channel Symbol Stream Processing, and (2) Decoding-Encoding Information Bit Stream Processing. They are described briefly below.

Hard Decision Channel Symbol Stream Processing. Each uplink transmitted waveform is optimally demodulated using a bank of matched filters. A hard decision is made as to the correct waveform identity. The decision is transmitted downlink with M-ary orthogonal modulation, as used on the uplink. The downlink demodulator consists of a bank of  $M$  matched filters with the vector of real valued outputs supplied to the decoder.



Decoding-Encoding Information Bit Stream Processing. The entire transmission is first demodulated and then decoded. The decoded data bits are then re-encoded, remodulated, and transmitted downlink. The satellite encoder and modulator are identical to the uplink encoder and modulator. The satellite decoder is identical to the downlink decoder. The satellite and downlink demodulators each consist of a bank of  $M$  matched filters with the vector of real valued outputs supplied to the decoders.

The parameter,  $R_0$ , having units of bits/uplink waveform, is used to quantify the end-to-end bit error performance of the system model with the two different forms of on-board processing described above. This parameter has become operationally much more important than channel capacity [Katz and Schnieder, 1978; Bordelon, 1977]. With any uplink and downlink decoding schemes, the end-to-end bit error probability will decrease with increasing  $R_0$ , although the actual functional variation may be quite complex. If the uplink encoder uses convolutional encoding, with the constraint length  $n$ , and the downlink decoder uses Viterbi decoding, then the end-to-end bit error probability is upper bounded by  $2^{-nR_0/R}$ , provided that  $R$ , the number of bits transmitted per uplink waveform, does not exceed  $R_0$ . If the uplink encoder uses convolutional encoding and the downlink decoder uses sequential decoding, then  $R_0$  provides an upper bound to the region of transmission rates (i.e., the largest  $R$ ) for which the computational requirements of sequential decoding are tolerable. An overview of the theoretical basis of  $R_0$  as a performance measure is given in appendix D.

A linear repeater (LRPT) satellite shares its available power proportionally between the uplink waveform and the uplink noise. Each is retransmitted with power proportional to that which it impresses on the repeater's transmitter. Thus, the degree of power robbing (by the noise) depends upon uplink conditions. With a repeater satellite used in the system model, the  $M$ -ary orthogonal signals transmitted uplink are retransmitted downlink and received at the ground in a white Gaussian noise environment. It is assumed that the optimum downlink demodulator is used. In this case, it consists of a bank of  $M$  matched filters. Each filter is matched to a different waveform in the uplink modulator signal set. The output of this optimum demodulator is the entire set of the (real valued) matched filter outputs.

With a linear repeater, the portion of the system extending from the uplink modulator to the downlink modulator can be

characterized as a discrete input-memoryless channel. The parameter  $R_0$  for this channel is shown to be:

$$R_0 \text{ (LRPT)} = -\log \left\{ \frac{1}{M} + \frac{M-1}{M} \exp \left[ -\frac{1}{2} \left( \frac{E_{sd}}{N_{od}} \right)_{\text{eff}} \right] \right\} \quad (3.38)$$

where

$$\left( \frac{E_{sd}}{N_{od}} \right)_{\text{eff}} = \frac{\frac{E_{su}}{N_{ou}}}{1 + \frac{C_u/N_{ou}}{C_d/N_{od}} + \frac{C_u/N_{ou}}{C_d/N_{od}} \frac{N_{ou} B}{C_u}}$$

All logarithms are to base 2.

Some discussion of eq. (3.38) is worthwhile. The power which the uplink noise presents to the transmitter of the linear repeater is  $N_{ou}B$  watts, and the power robbed is represented by the fraction of total transmitter power allocated to the noise. This fraction is given by:

$$F_{\text{pr}} = \frac{N_{ou} B}{C_u + N_{ou} B} \quad (3.39)$$

The performance degrading effect of this noise power robbing is evident in eq. (3.38) through the exponent denominator terms. In the limit when

$$\frac{C_u/N_{ou}}{C_d/N_{od}} \rightarrow 0$$

(perfect downlink), eq. (3.38) becomes

$$R_0 \text{ (LRPT)} \quad R^\infty = -\log \left[ \frac{1}{M} + \frac{M-1}{M} \exp \left( -\frac{1}{2} \frac{E_{su}}{N_{ou}} \right) \right] \quad (3.40)$$

$R_0^\infty$  represents the optimum end-to-end performance that can be obtained with any type of satellite processing used in the figure 3-25 model. It corresponds to the  $R_0$  achieved when the uplink waveforms are optimally demodulated on board the satellite and then retransmitted to the ground on a perfect downlink.

Figures 3-30, 3-31, and 3-32 are derived from the  $R_0$  (LRPT) expression given by eq. (3.38) and include the degradation for each part of the link given in table 3-3. These figures illustrate the variation of the  $R_0$  (LRPT) as a function of the parameter pairs ( $E_{su}/N_{ou}$ ,  $E_{sd}/N_{od}$ ) with  $M = 4$ . Similar figures are derivable for other values of  $M$ .

In the following discussion of the two forms of on-board signal processing, their performance is compared to that of the linear repeater with degradation added to account for non-linear elements.

Hard Decision Processing (HDP). With HDP, uplink waveforms only are demodulated on board the satellite. No decoding takes place. The demodulator output is, itself, remodulated and transmitted downlink. Consequently, the portion of the system model extending from the uplink modulator to the downlink demodulator can be characterized as a tandem connection of two, discrete input, memoryless channels.

This tandem connection of two channels is itself a discrete input, memoryless channel. Hence, with HPD the parameter  $R_0$  can be used to describe the end-to-end link performance of the system model. The expression for  $R_0$  is easily obtained from [Massey, 1974] since the end-to-end channel transition probabilities can be determined from the respective uplink and downlink channels.  $R_0$ (HDP) for this channel is given by:

$$R_0 \text{ (HDP)} = -\log \sum_i^M \left[ \frac{1}{M} \sum_j^M \sqrt{P_{ij}} \right]^2 \quad (3.41)$$

where  $P_{ij}$  is the transition probability between the  $j$ th input and the  $i$ th output letter of  $M$ -ary alphabet. A bound to eq. (3.41) is given in [Katz, Schneider, 1977] and is more convenient to use for large  $M$ .

$R_0$  (HDP) is shown in figures 3-33, 3-34, and 3-35 for the three backoff states in table 3-3. It can be seen in comparing  $R_0$  of the linear repeater and HDP that  $R_0$  (HDP) falls off sharply for low-

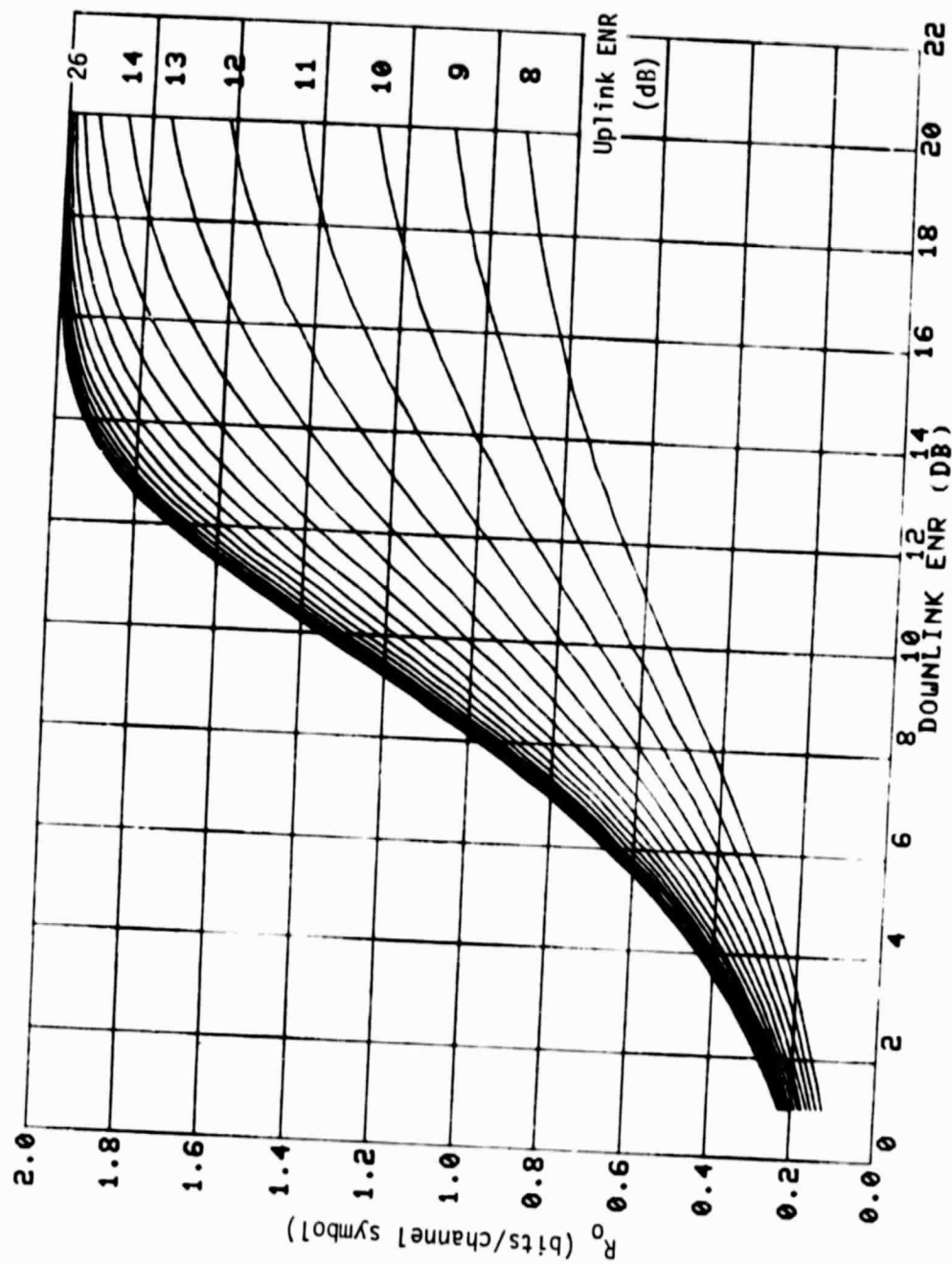


Figure 3-30.  $R_o$  for Conventional (LRPT) Repeater with HPA Backoff = 6 dB,  
TWTA Backoff = 4 dB

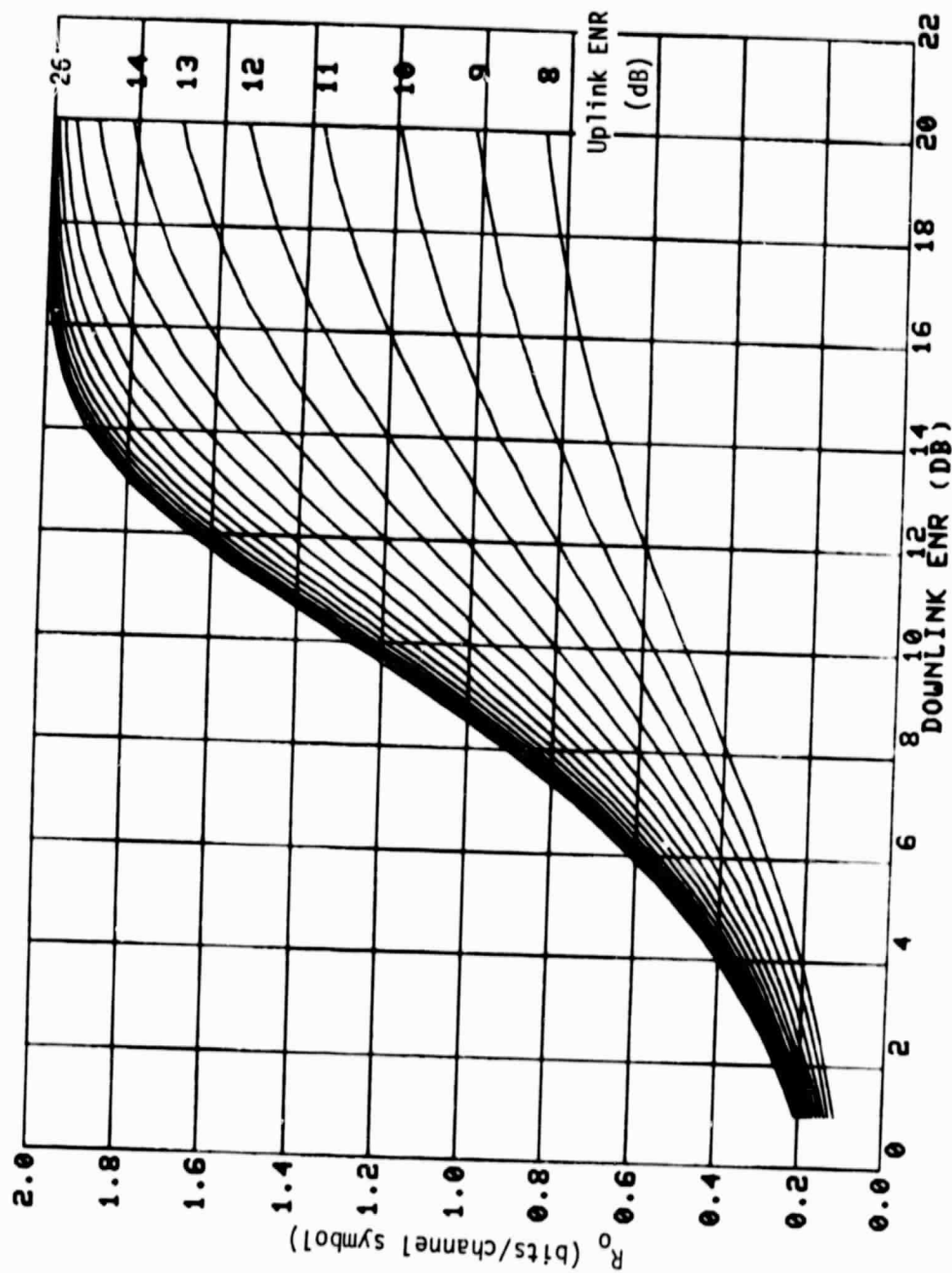


Figure 3-31.  $R_o$  for Conventional (LRPT) Repeater with HPA Backoff = 6 dB,  
TWTA Backoff = 0 dB

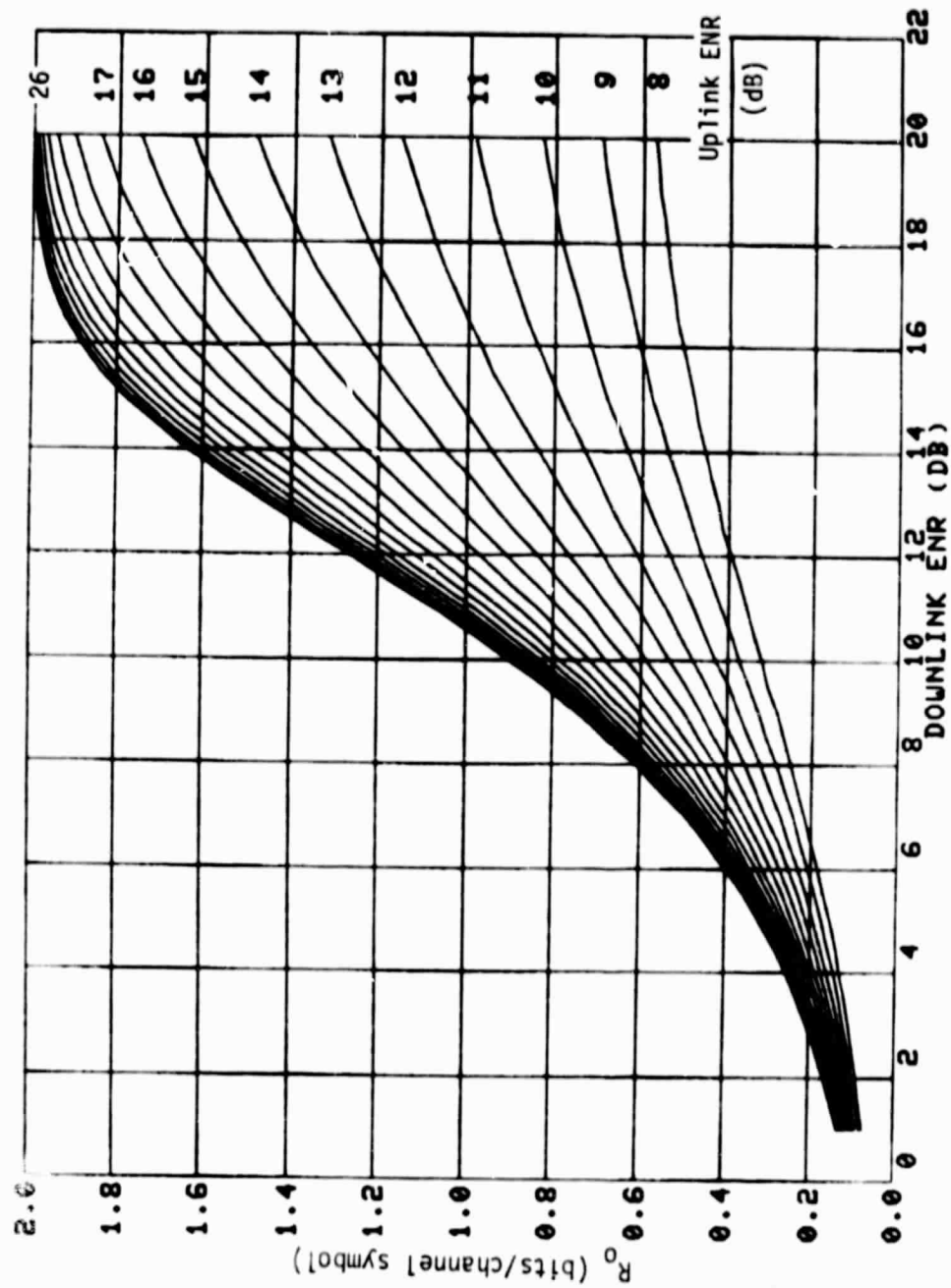


Figure 3-32.  $R_0$  for Conventional (LRPT) Repeater with HPA Backoff = 0 dB,  
TWTA Backoff = 0 dB

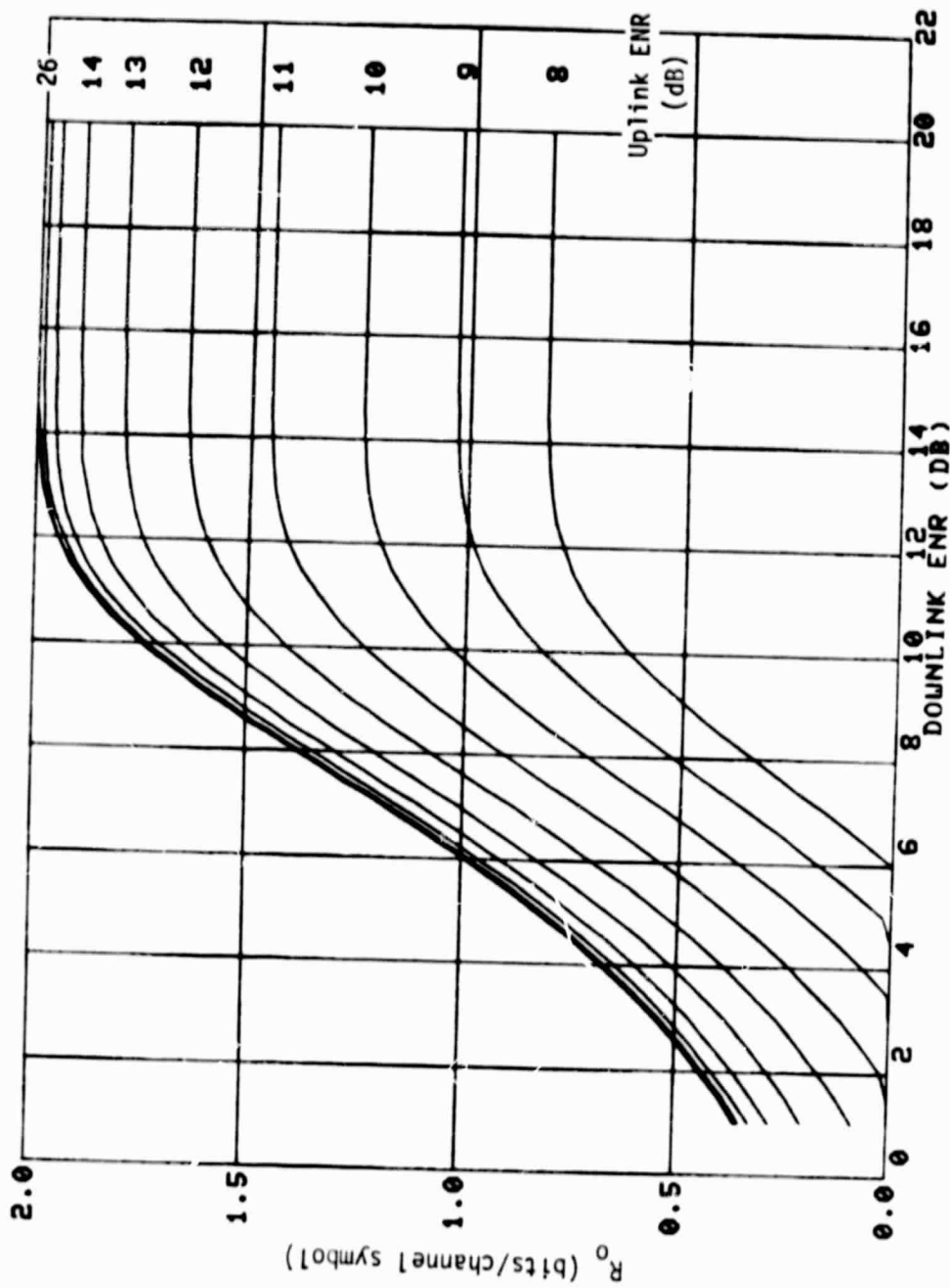


Figure 3-33.  $R_0$  for HDP with HPA Backoff = 6 dB, TWTA Backoff = 4 dB

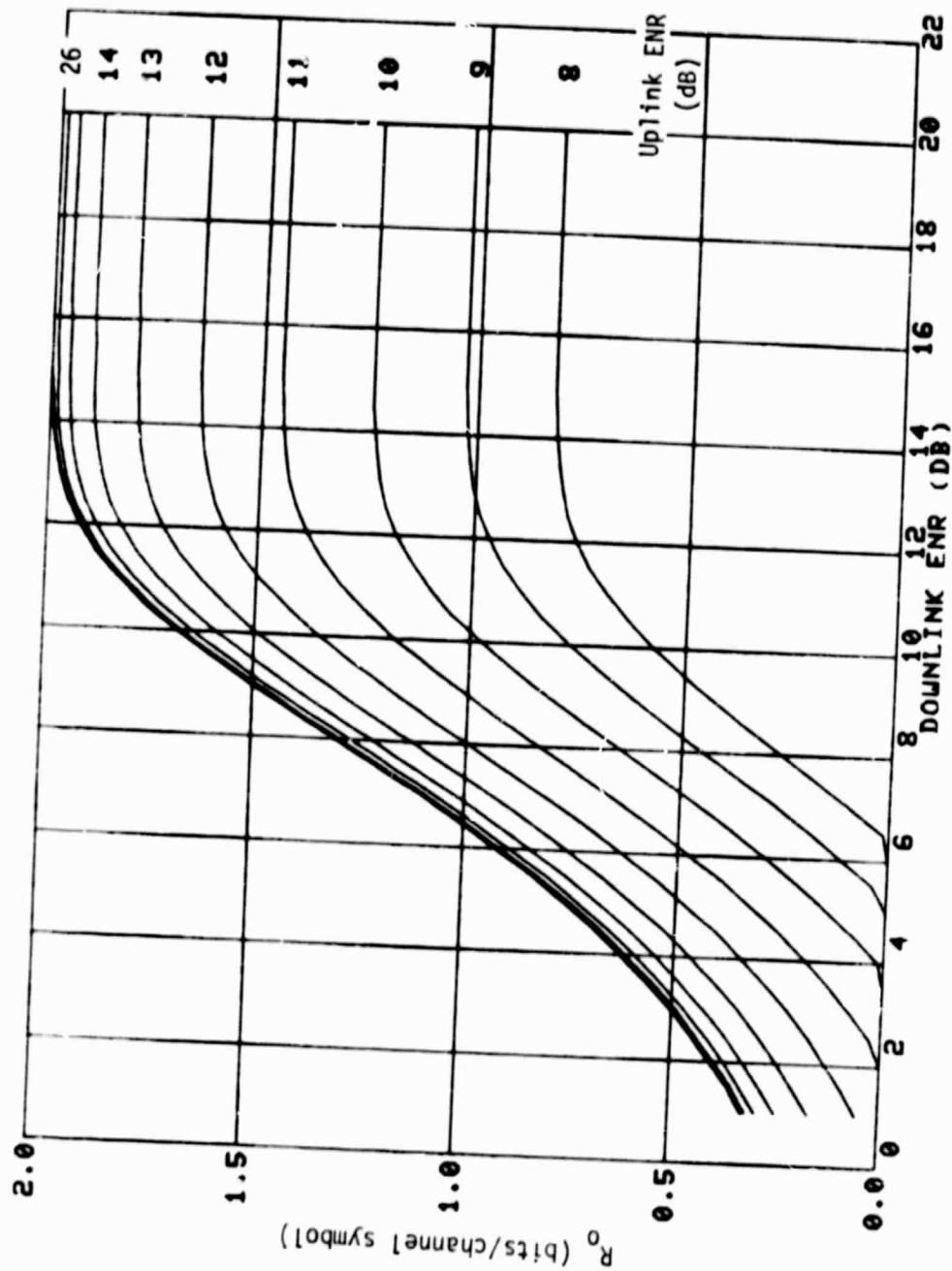


Figure 3-34.  $R_0$  for HPD with HPA Backoff = 6 dB, TWT Backoff = 0 dB



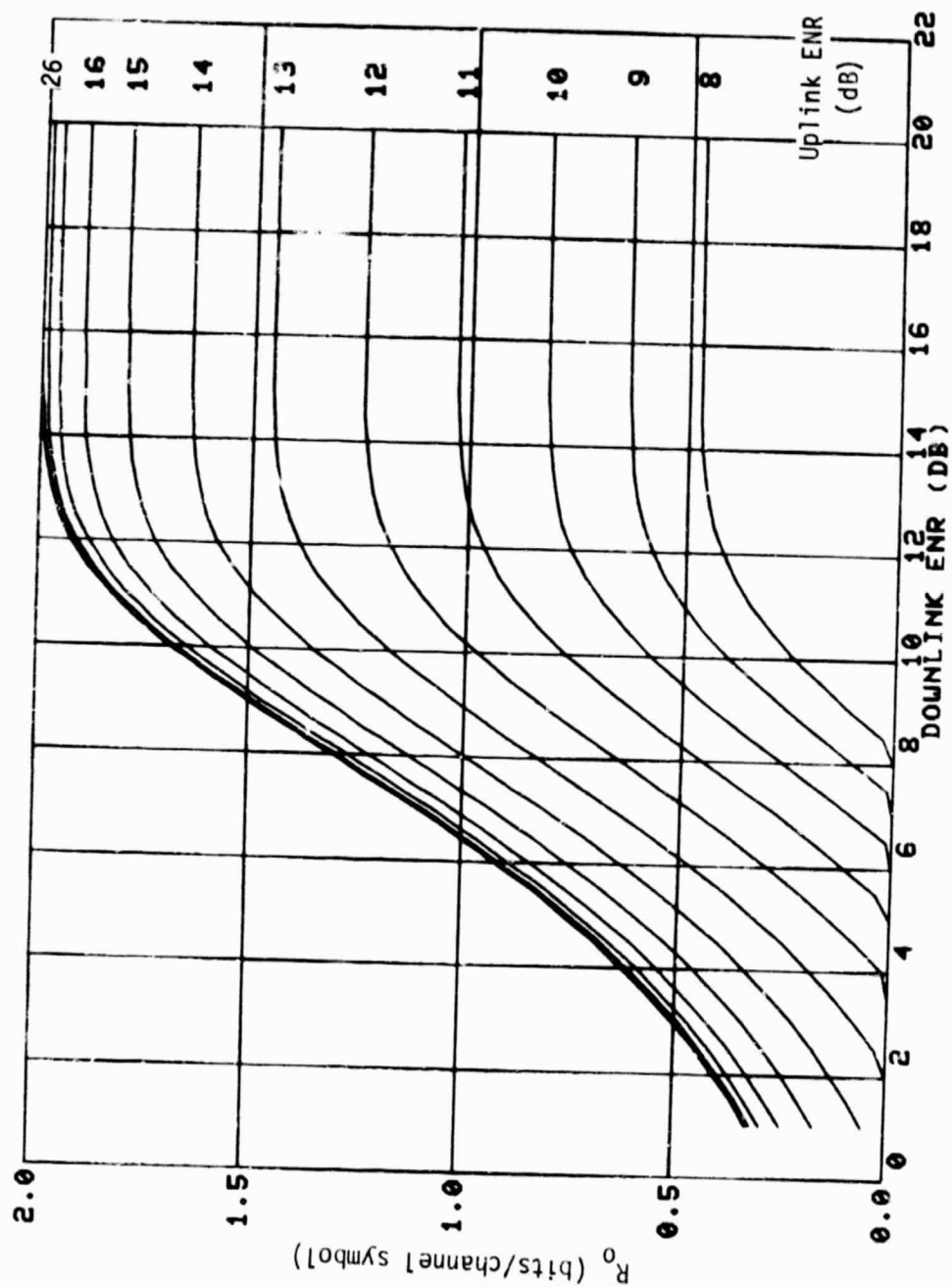


Figure 3-35.  $R_0$  for HDP with HPA Backoff = 0 dB, TWTA Backoff = 0 dB

uplink  $E_{su}/N_{ou}$ . This is to be expected since some statistical information about the uplink signal is lost by the hard decision. For low-uplink  $E_{su}/N_{ou}$  this dominates the HDP performance.

We can conclude from the foregoing results and the hypothesized implementation that HDP has the following properties. The uplink bandwidth requirement is the same as for the linear repeater; however, the downlink bandwidth requirement may be less than that needed for the linear repeater. No on-board storage is required by this processor. The degrading effect of noise power robbing (present in the linear repeater) is eliminated. A disadvantage, however, is that the hard decision demodulation denies the ground receiver (downlink demodulator and decoder) some statistical information about the uplink transmission process. With HDP, the ground receiver has no information about the absolute values of the satellite demodulator matched filter outputs. It knows nothing about the closeness of the decision and, as a result, the linear repeater can perform better with HPD when uplink  $E_{su}/N_{ou}$  is low. However, if we include the degradation for the link indicated in table 3-3, HDP always exceeds the performance of the linear repeater.

Decoding-Encoding Processing (DEP). With DEP, the uplink and downlink transmission can be characterized as two complete communications systems connected in tandem. Each received uplink waveform is demodulated on board the satellite with a bank of  $M$  matched filters. Each filter is matched to a different signal in the set employed by the uplink modulator. The  $M$  real valued matched filter outputs are supplied to a decoder. The decoded digits are re-encoded, modulated, and transmitted downlink. Although they need not be, the uplink encoder and satellite encoder are assumed identical and the satellite modulator is assumed to see  $M$ -ary orthogonal signaling. The downlink demodulator consists of a bank of  $M$  matched filters each matched to a different waveform in the satellite modulator signal set. The real value matched filter outputs are supplied to the downlink decoder which is assumed identical to the satellite decoder. Both the uplink and downlink transmission are considered coherent. The expression for  $R_0$  (DEP) for this tandem channel is given in eq. (3.42)

$$R_0 \text{ (DEP)} \geq \min (R_u, R_d) \quad (3.42)$$

where

$$R_u = 1 - \log_2 \left[ 1 + \exp \left( - \frac{1}{2} \frac{E_{su}}{N_{ou}} \right) \right]$$

$$R_d = 1 - \log_2 \left[ 1 + \exp \left( - \frac{1}{2} \frac{E_{sd}}{N_{od}} \right) \right]$$

or some highly quantized version.

Plots of the lower bound to  $R_0$  (DEP) are shown in figures 3-36, 3-37, and 3-38 for the three backoff states in table 3-3. When we compare performance, DEP easily exceeds the performance of the linear repeater, or HDP, and achieves this performance without requiring excess bandwidth. DEP achieves its greatest difference in a region of relative low uplink  $E_{su}/N_{ou}$ . This attribute alone makes DEP a candidate for satellite systems which need to work with small, low-powered ground terminals.

As with HDP, DEP eliminates the possibility of noise power robbing at the satellite. With the version of DEP being considered, the downlink bandwidth requirement is the same as that of HDP, B Hz. In contrast to HDP, DEP employs all the available information about the uplink transmission process. No information is destroyed with the on-board processing. Information is not under-utilized.

#### 3.4.4 Analysis of Results

Results for a regenerative repeater with coded data can be analyzed using the same approach as that for uncoded data. Curves with constant  $R_0$  as a function  $E_u/N_{ou}$  and  $E_d/N_{od}$  are plotted and shown in figures 3-39, 3-40, and 3-41 for each type of satellite channel. For a channel with a given  $R_0$ , the information bit error rate bound can be determined, and is a function of code complexity and information rate  $R$  where  $R < R_0$ . By using  $R_0$  as a performance measure, the specific code need not be considered.

As in the case of uncoded data, the points tangent to a -1 slope on each curve are used to measure the difference between a conventional repeater and a regenerative repeater. Such points represent the joint minimization of the uplink and downlink  $E/N_0$  for  $R_0 = 1.4$ . Results are summarized in table 3-5.

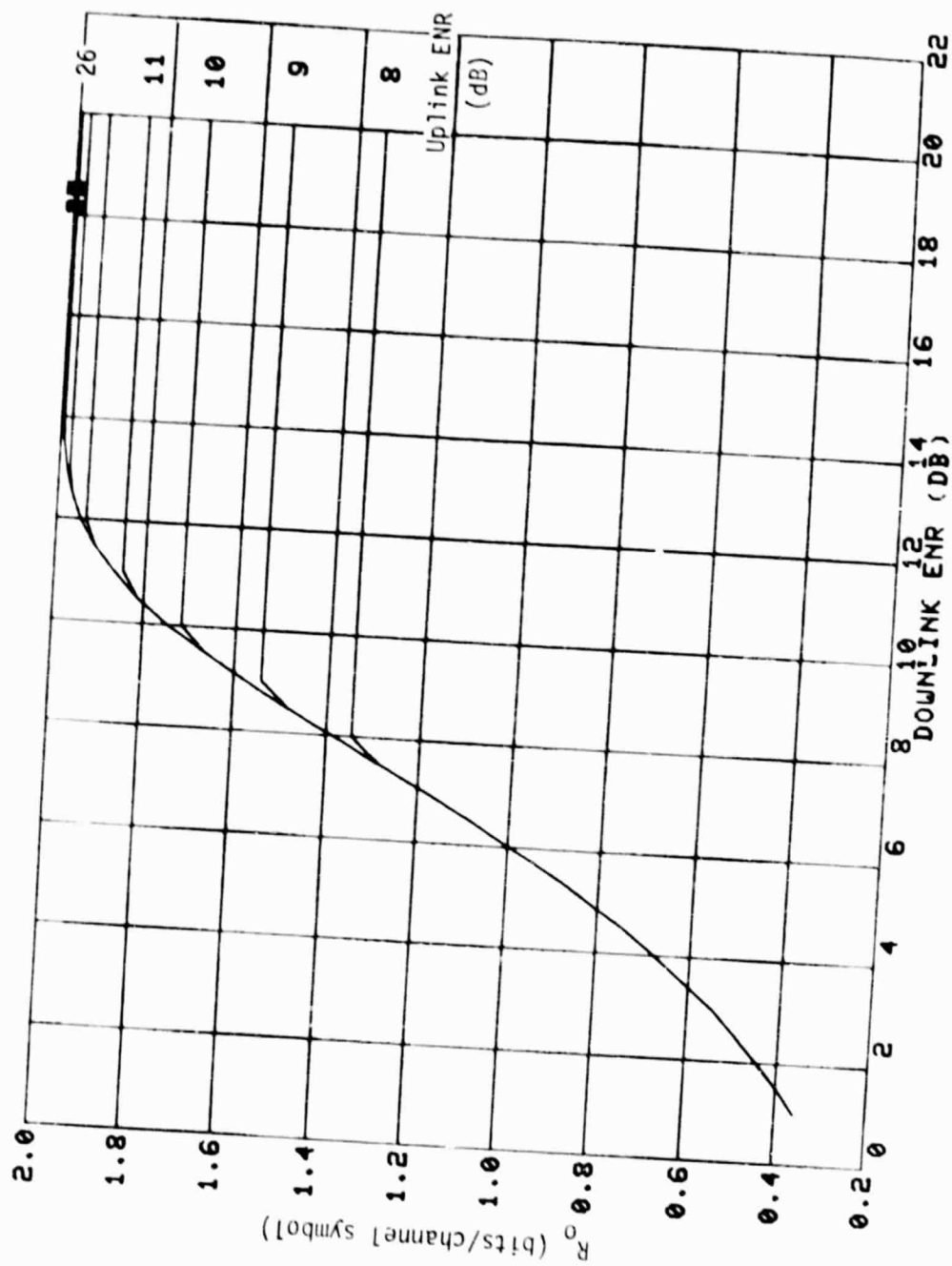


Figure 3-36.  $R_0$  for DEP with HPA Backoff = 6 dB, TWTA Backoff = 4 dB

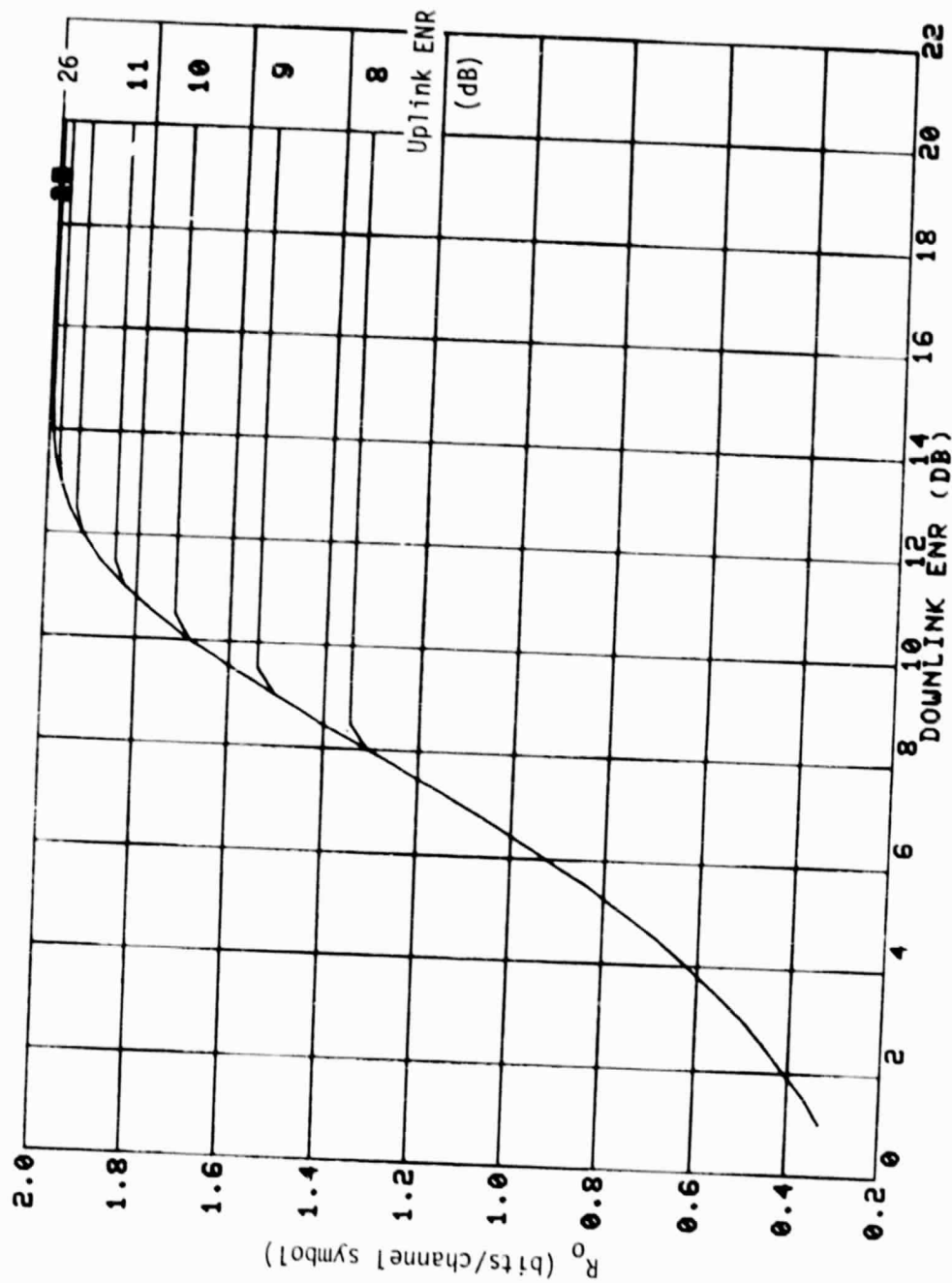


Figure 3-37.  $R_0$  for DEP with HPA Backoff = 6 dB, TWTA Backoff = 0 dB

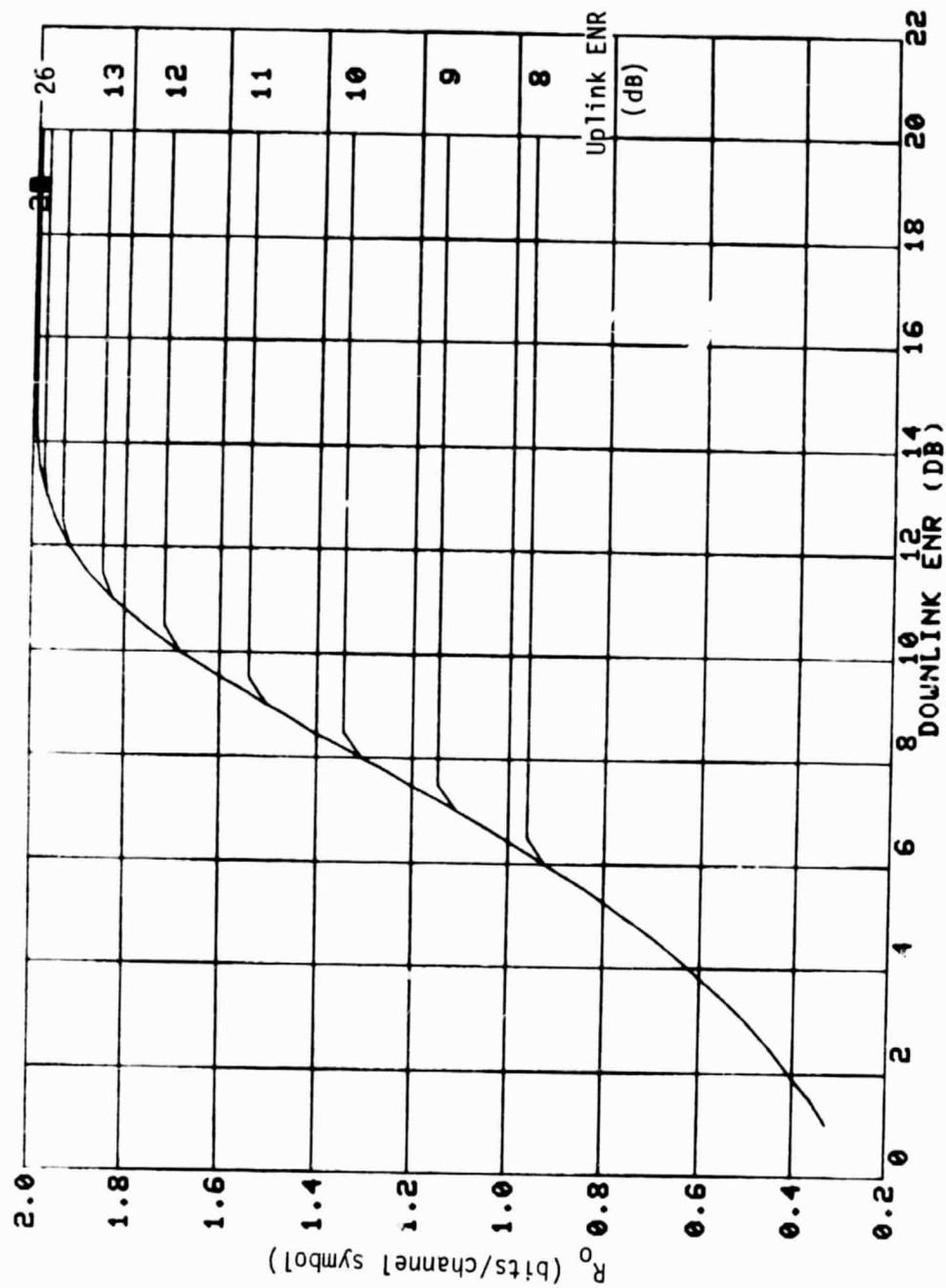


Figure 3-38.  $R_0$  for DEP with HPA Backoff = 0 dB, TWTA Backoff = 0 dB

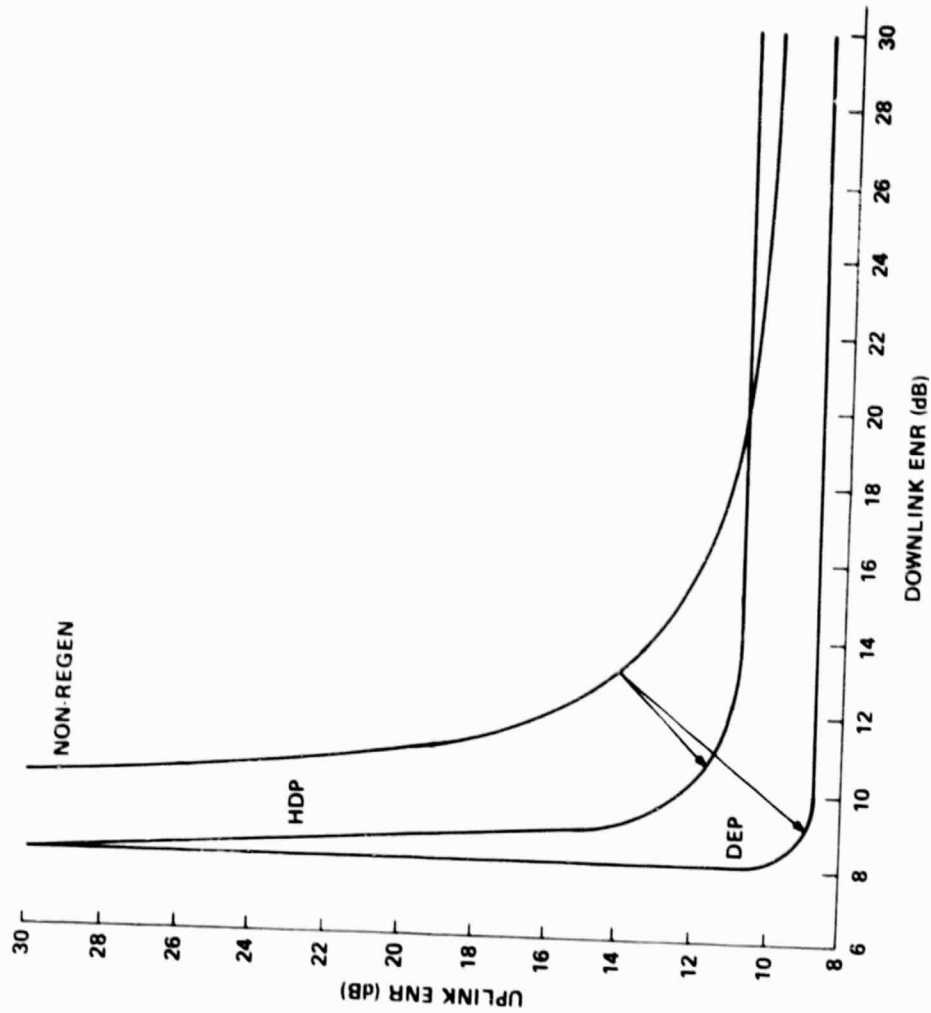


Figure 3-39. Performance Curves for Constant  $R_0$  of 1.4 Bits/Channel Symbol  
With HPA Backoff = 6 dB, TWT Backoff = 4 dB

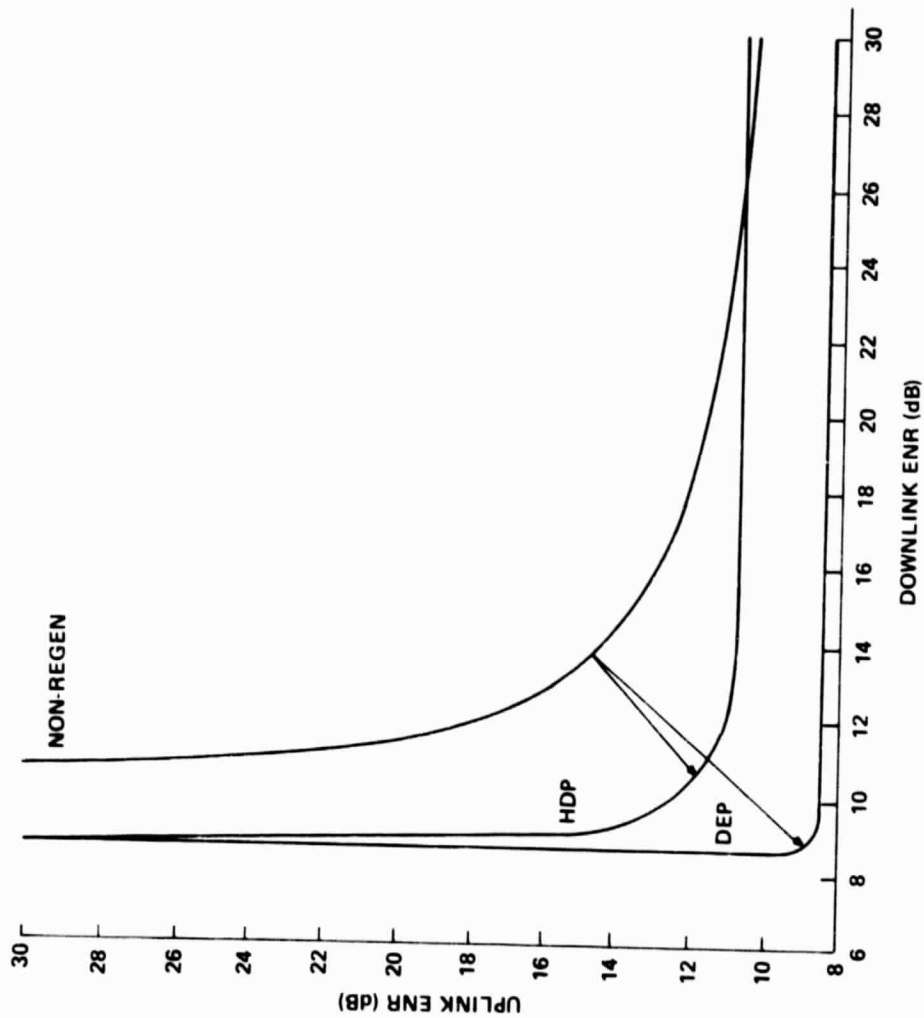


Figure 3-40. Performance Curves for Constant  $R_0$  of 1.4 Bits/Channel Symbol  
With HPA Backoff = 6 dB, TWTA Backoff = 0 dB



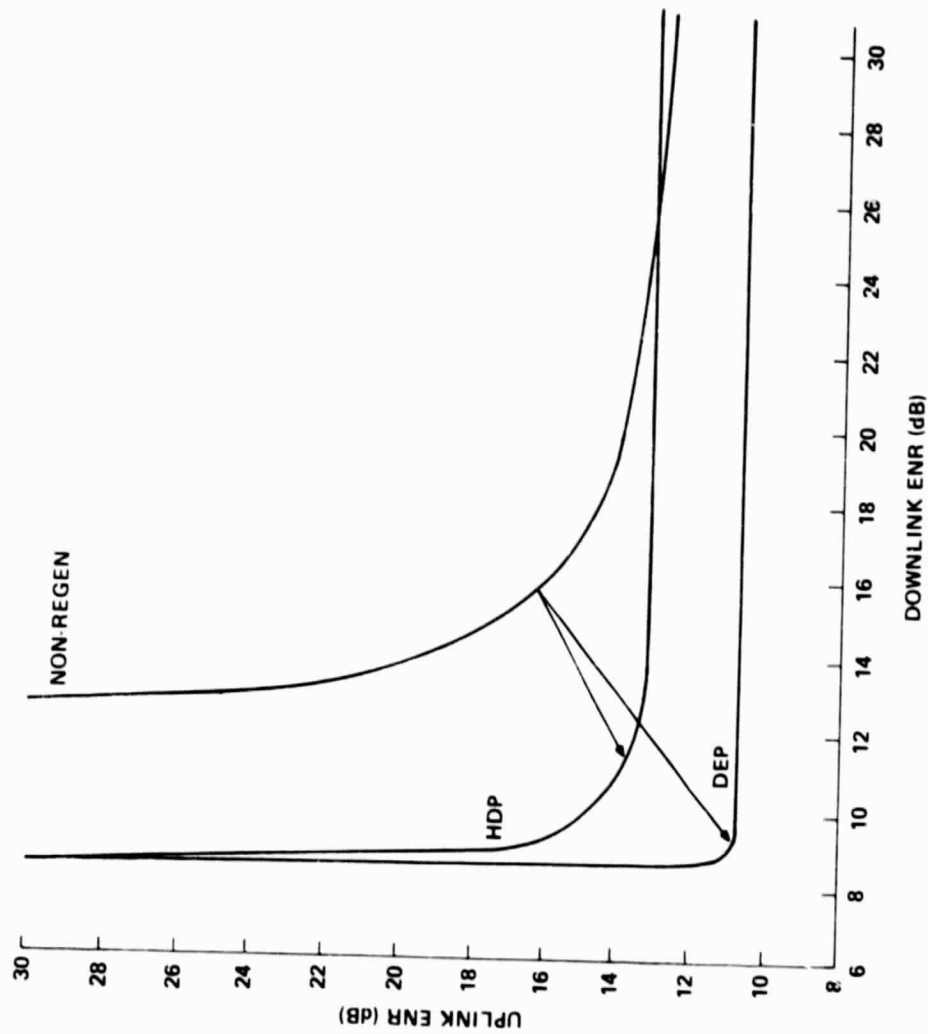


Figure 3-4]. Performance Curves for Constant  $R_0$  of 1.4 Bits/Channel Symbol  
With HPA Backoff = 0 dB, TWTA Backoff = 0 dB

Table 3-5

Difference in Conventional and Regenerative Link  $E/N_0$   
Values for Uplink and Downlinks

HPA Output Backoff (dB)	Satellite TWT Backoff (dB)	Uplink Difference		Downlink Difference	
		HDP (dB)	DEP (dB)	HDP (dB)	DEP (dB)
0	0	3.0	4.0	6.5	5.5
6	0	2.75	3.0	5.9	4.9
6	4	2.4	3.0	2.4	4.0

It should be noted that an operating point designated by any pair of uplink and downlink energy-to-noise ratios (ENRs) which lie above a respective curve in figures 3-27, 3-28, 3-29 for uncoded data or figures 3-39, 3-40, and 3-41 for coded data will result in a lower BER or increased value of  $R_0$ . For all backoff cases examined, the operating points for a given performance level are significantly lower for the regenerative repeater case than for the conventional repeater case. That is, the regenerative repeater results in either a greater fade margin or a reduced power requirement on both up and downlinks. It should also be noted that the improvement potential is greatest for the 0/0 HPA/TWT backoff case in which the degradation is also the greatest. This indicates that the regenerative repeater accomplishes its greatest improvement when the degradation encountered for the conventional case is the greatest.

Although the above discussion compares performance on the basis of minimum uplink and downlink ENRs, it is clear that other operating points can be selected to achieve other objectives. For example, from figure 3-41 we see that if an uplink ENR of 14 dB is selected, the resulting improvement in downlink ENR between HDP and nonregeneration is 8.5 dB. That is, the downlink ENR can be minimized at the expense of the uplink and still achieve the same performance.

## REFERENCES

Abramson, N., "The Throughput of Packet Broadcasting Channels," IEEE Transactions on Communications, Vol. COM-25, No. 1, January 1977, pp. 117-128.

Acampora, A. S., and Langseth, R. E., "A 300 Megabaud Frame Synchronizer and Interface Design for Variable Capacity TDMA Ground Terminals," scheduled for IEEE Intern. Conf. Commun., Boston, MA, 10-14 June 1979.

Aulin, T., Rydbeck, N., and Sundberg, C.-E., "Bandwidth Efficient Digital FM With Coherent Phase Tree Demodulation," scheduled for IEEE Intern. Conf. Commun., Boston, MA, 10-14 June 1979.

Aulin, T., and Sundberg, C.E., "M-ary CPFSK Type of Signaling With Input Data Symbol Pulse Shaping - Minimum Distance and Spectrum," scheduled for IEEE Intern. Conf. Commun., Boston, MA, 10-14 June 1979.

Bekey, I., "Preliminary Definition and Evaluation of Advanced Space Concepts, Volume I: Executive Summary, Volume II: Analyses and Results," ATR-78(7674)-1, The Aerospace Corporation, 30 June 1978.

Bohacek, P. K., "The Implementation of CCIS In The Bell System," National Telecommunications Conference Record, Alabama, December 1978, pp. 31.1.1-31.1.5.

Bordelon, D. L., "On Signal Design by the  $R_0$  Criterion for Non-White Gaussian Noise Channels," NTC 77, December 1977.

Campanella, S. J., Assal, F., and Berman, A., "On-Board Regenerative Repeaters," ICC 77, June 1977.

CCITT Blue Book, Vol. II, Geneva, 1965, p. 239.

Cooper, R. B., Introduction to Queueing Theory, New York: The McMillan Company, 1972.

Cover, T. M., "Broadcast Channels," IEEE Trans. Inform. Th., Vol. IT-18, January 1972, pp. 2-14.

Dahlbom, C. A., and Ryan, J. S., "History and Description of a New Signaling System," The Bell System Technical Journal, Common Channel Interoffice Signaling, February 1978, Vol. 57, No. 2, pp. 225-250.

de Jager, F., and Dekker, C. B., "Tamed Frequency Modulation, A Novel Method to Achieve Spectrum Economy in Digital Transmission," IEEE Trans. Commun., Vol. COM-26, May 1978, pp. 534-542.

Eaves, R. E., and Wheatley, S. M., "Optimization of Quadrature-Carrier Modulation for Low Crosstalk and Close Packing of Users," IEEE Trans. Commun., Vol. COM-27, January 1979, pp. 176-185.

Edelson, B. I., and Morgan, W. L., "Orbital Antenna Farms," Astronautics and Aeronautics, September 1977, pp. 20-29.

Fordyce, S. W., and Jaffe, L., "Future Communications Concepts: The Switchboard-in-the-Sky, Part I: The Problem; Part II: A Solution," Satellite Communications, Vol. 2, February 1978, pp. 22-26, March 1978, pp. 22-27.

Forney, G. D., Jr., "The Viterbi Algorithm," Proc. IEEE, Vol. 61, March 1973, pp. 268-278.

Frerking, R. F., "Enhanced Signaling for New Customer Services and Network Features," National Telecommunications Conference Record, Alabama, December 1978, pp. 31.2.1-31.2.3.

Gerla, M., and Kleinrock, L., "On the Topological Design of Distributed Computer Networks," IEEE Transactions on Communications, Vol. COM-25, No. 1, January 1977, pp. 48-60.

Gossen, R. N., Jr., "100,000+ Gates on a Chip: Mastering the Minutia," IEEE Spectrum, Vol. 16, March 1979, pp. 42-47.

Hilborn, C. G., et al., "Implications of Demand Assignment for Future Satellite Communication Systems," Final Report, DCA 100-76-C-0060, System Control Inc., Palo Alto, June 1977.

Jain, P., "Utilization of EHF by Future Military COMSATS," scheduled for IEEE Intern. Conf. Commun., Boston, MA, 10-14 June 1979.

Kalet, I., and White, B. E., "Suboptimal Continuous Shift Keyed (CSK) Demodulation for the Efficient Implementation of Low Crosstalk Data Communication," IEEE Trans. Commun., Vol. COM-25, September 1977, pp. 1037-41.

Katz, J. L., and Schneider, K. S., "On-Board Regeneration Processor for Digital Communication Satellites," AIAA 7th Communication Satellite Systems Conference, April 1978.

Kleinrock, L., Communication Nets: Stochastic Message Flow and Delay, New York: Dover Publications, Inc., 1964.

Kleinrock, L., Queueing Systems, Vol. I: Theory, New York: Wiley-Interscience, 1976.

Kleinrock, L., and Scholl, M., "Packet Switching in Radio Channels: New Conflict-Free Multiple Access Schemes for a Small Number of Data Users," Proceedings of International Conference on Communications, Chicago, Illinois, June 1977, pp. 22.1-105 to 22.1-11).

Kota, S. L., "Survey of Demand Assignment Multiple Access (DAMA) Techniques for Satellite Communications," M-78-239, The MITRE Corporation, January 1979.

Lusignan, B. B., "Single-sideband Transmission for Land Mobile Radio," IEEE Spectrum, Vol. 15, July 1978, pp. 33-37.

Martin, J., Telecommunications and the Computer, Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1976.

Massey, J. L., "Coding and Modulation in Digital Communications," Proceedings, International Zurich Seminar on Digital Communications, Zurich Switzerland, March 12-15, 1974, pp. E2(1) - E2(4).

McGarty, T. P., and Warner, T. H., "Multiple Beam Satellite System Optimization," IEEE Trans. Aerospace Elect. Systems, Vol. AES-13, September 1977, pp. 504-511.

Metzger, L. S., "Performance of Phase Comparison Sinusoidal Frequency Shift Keying," IEEE Trans. Commun., Vol. COM-26, August 1978, pp. 1250-1253.

Mier, M. G., Blasingame, J. M., Johnson, R. L., Cummins, S. E., and Buvinger, E. A., "Magnetic Bubble Memories for Military Applications: 1978 and Beyond," presented at 6th IEEE Texas Conf. Computing Systems, 15 November 1977, Austin, TX.

Morrow, W. E., Jr., "Current and Future Communication Satellite Technology, Proc. XXVI Intern. Astronautical Congress, ed. L. G. Napolitano, Pergamon Press, New York, 1976, pp. 277-296.

Nowack, H. E., "A Prospectus for the Evolution of the CCIS Network," National Telecommunications Conference Record, Alabama, December 1978, pp. 31.4.1-31.4.4.

Reiffen, B. and White, B. E., "On Low Crosstalk Data Communication and Its Realization by Continuous-Frequency Modulation Schemes," IEEE Trans. Commun., Vol. COM-26, January 1978, pp. 131-135.

Reudink, D. O., "Spot Beams Promise Satellite Communication Breakthrough," IEEE Spectrum, Vol. 15, September 1978, pp. 36-42.

Rydbeck, N., and Sundberg, C.-E., "Recent Results on Spectrally Efficient Constant Envelope Digital Modulation Methods," scheduled for IEEE Intern. Conf. Commun., Boston, MA, 10-14 June 1979.

Staelin, D. H., "Issues in the Establishment of a Large National Broadband Switched Communications Network," unpublished work presented to Boston Section IEEE Commun. Soc. and Inform. Th. Chapter, 29 November 1978.

White, B. E., "A Worst-Case Crosstalk Comparison Among Several Modulation Schemes," IEEE Trans. Commun., Vol. COM-25, September 1977, pp. 1032-1037.

White, B. E., Kalet, I., and Heggstad, H. M., "Offset Quadrature Phase-Comparison Modulation Schemes for Low-Crosstalk Communication," Proc. IEEE Intern. Conf. Commun., 13-15 June 1977, Chicago, IL, pp. 6.5-133 through 6.5-137.

Wolf, J. K., "Multiple Access Communications Using M-ary FSK," unpublished work presented to Boston Section IEEE Comm. Soc. and Inform. Th. Chapter, 14 February 1979.

## SECTION 4

### BASEBAND PROCESSOR IMPLEMENTATION CONSIDERATIONS

Candidate satellite designs for obtaining the desired connectivity and data throughput considered feasible in the 1990 to 2000 time frame have been presented. This section deals with factors affecting the digital hardware implementation of the baseband processing section of the satellite, as shown within the dotted lines of figure 4-1. The concepts considered assume capabilities beyond the current state-of-the-art for digital hardware components and processor architectures.

The region between the multiple uplink demodulator inputs and the multiple downlink modulator outputs is of major interest (see figure 4-2). Areas not covered in this section include global systems timing concepts, the required circuit redundancy to obtain lifetime reliability goals, overall system timing and link synchronization factors, and built-in test equipment algorithms that may prove necessary to monitor the performance of the processing hardware.

#### 4.1 ASSUMPTIONS

The baseband digital processing portions of the satellite can be modeled as a computer directing an array of peripheral processors which handle the multiple processing tasks. Multiple memories are partitioned and sized to handle data storage requirements encountered for moderate system delays to support the connectivity requirements. Before hardware implementation is examined, the following interface assumptions are presented.

##### Uplink Format

- Uplink beams are comprised of fixed and scanning beams.
- Each beam handles up to N FDMA uplinks.
- For the fixed uplink beams, the average data rate per FDMA channel will be either T1 (1.544 Mb/s) or T2 (6.312 Mb/s) plus appropriate overhead bits. (T2 channels are basically comprised of four T1 channels with slight format changes.)

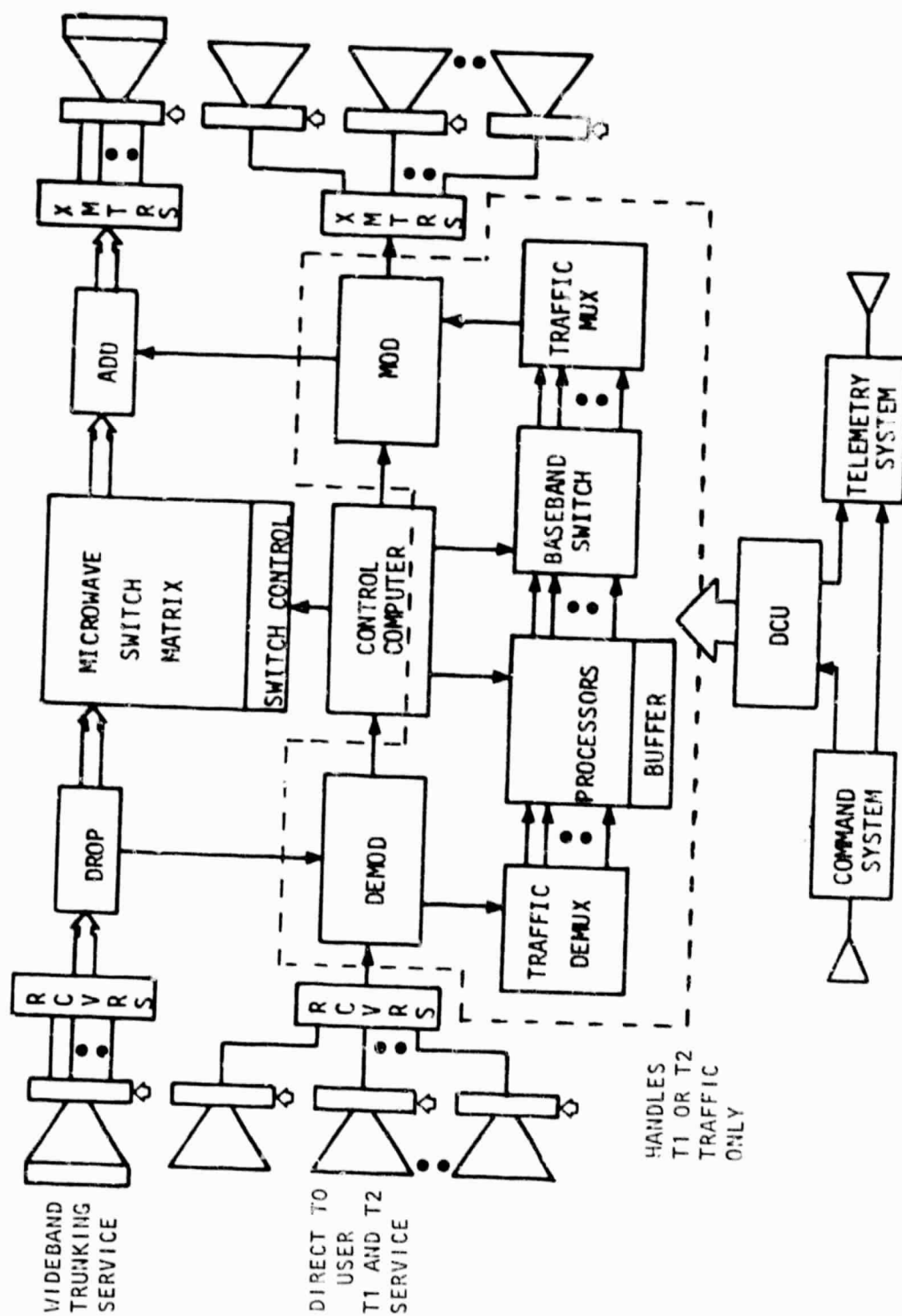


Figure 4-1. Processing Satellite System



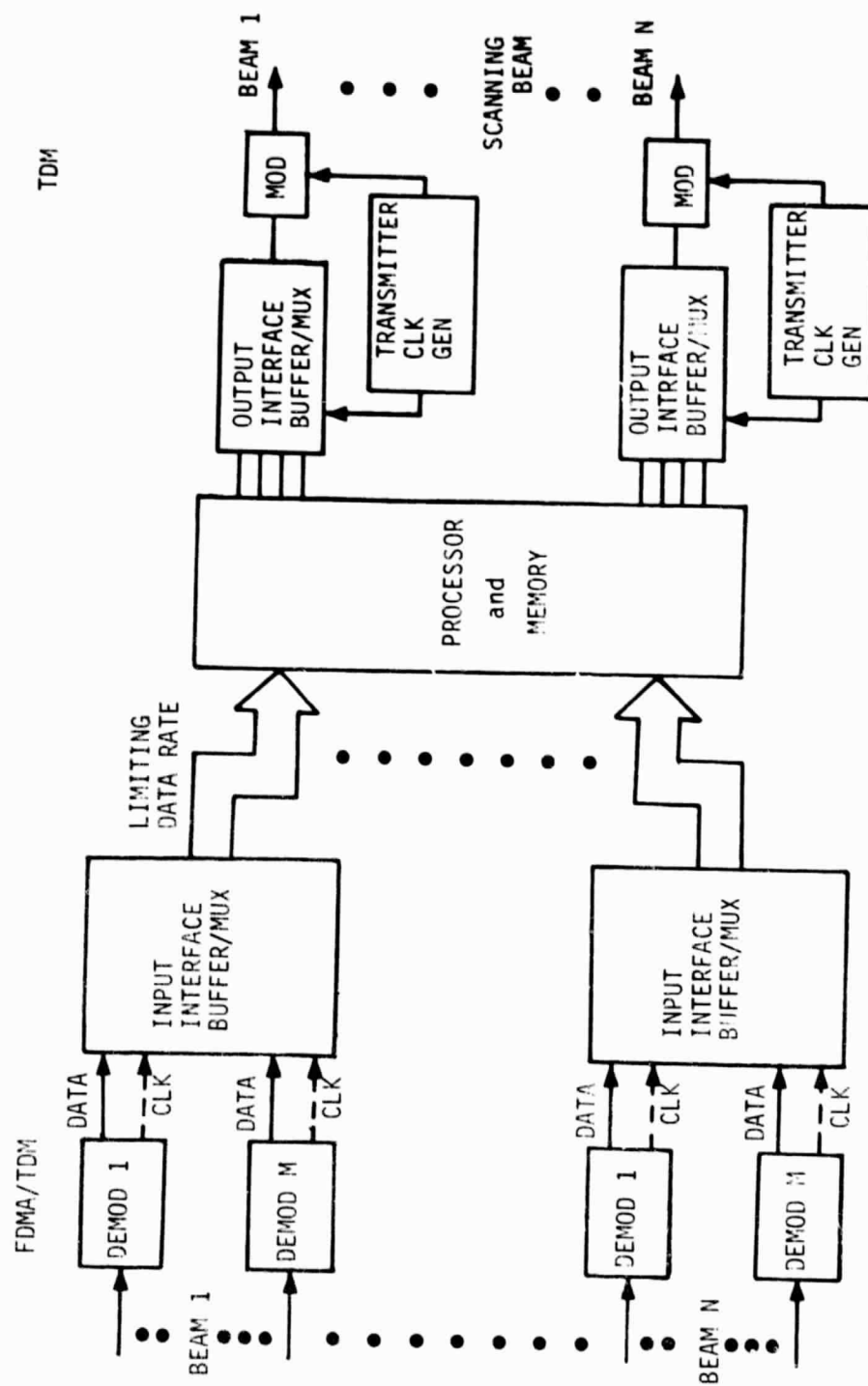


Figure 4-2. FDMA/TDM Processor Configuration

- For the scanning beams, data rates per FDMA channel will depend on the number of areas covered in a particular scan frame period. In a T2 channel, if the earth scan frame period is 100 msec and the beam dwell time is 10 msec, then the channel burst rate would be at least 10 times the T2 rate or  $>63.12 \text{ Mb/s}$ .

#### Downlink Format

- N downlink beams are comprised of fixed and scanning beams.
- Very high burst rate TDM techniques are used on each beam.
- Burst rates for the scanning beams are determined by the scan pattern of the beam.

### 4.2 BASEBAND PROCESSOR IMPLEMENTATION AND TECHNOLOGY

The conceptual baseband processor design illustrated in figure 4-3 is used as a vehicle for relating communication requirements to implementation and technology issues. The discussion here presents an overview of the processor operation from a signal flow view point, including a detailed description of the processor system elements.

Signals from each uplink beam are assumed to be pre-processed so that individual FDMA channels are present at IF to the digital processor. This is the assumption made here although there are techniques, such as digital Fourier transform (DFT), in which the FDMA filtering function and demodulation are performed in one operation. The tradeoffs between the various techniques are well understood and the ultimate decision as to which approach is best for the system considered here remains to be determined.

#### 4.2.1 Demodulation Processor Module

The pre-processed FDMA signals from each beam serve as the input to the demodulation processor module. This module contains M demodulators; one for each FDMA channel. The data demodulators are conceived as essentially all-digital programmable units whose input IF signal is switch-selectable from a subset of individual FDMA channel downconverters by the master controller. All uplink demodulators have identical capabilities and are interchangeable.

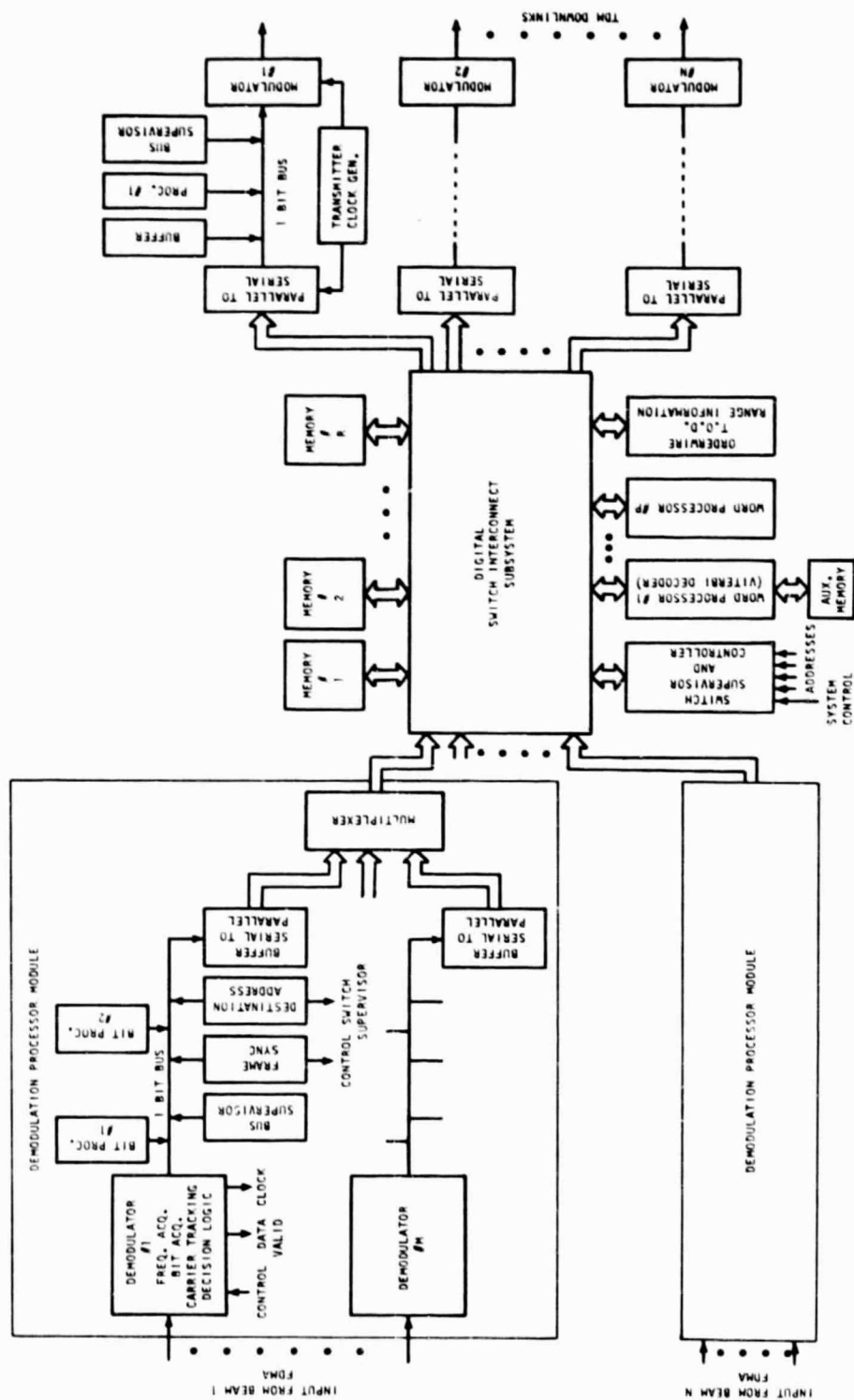


Figure 4-3. Baseband Processing System

Bit timing acquisition time goals of 1 to 2 bit intervals are assumed to minimize overhead.

Other inputs are:

- Operating rate control: at  $T_1$ ,  $K \times T_1$ ,  $T_2$ , or  $K \times T_2$ , rates are minimum where  $K$  is the scanning beam frame period/dwell time ratio.
- Frame sync and burst sync word inputs, for synchronizing the start of actual data reception, are loaded once per frame and stored in holding registers. High speed multi-bit digital comparators are used to perform detection tasks.

Outputs include:

- A start-of-data or data valid flag signal marking the data portion of each burst, which enables clocking of the data into the input serial-to-parallel register.
- Bursts of synchronous clock and data are coincident with the above flag. Rates are  $T_1$ ,  $K \times T_1$ ,  $T_2$ , or  $K \times T_2$  b/s.

The output of the demodulators feeds a one bit bus on which hang additional bit and function processors. bit processors are programmable and perform specific operations such as decryption or de-interleaving of the bit stream. Other special functions required by a specific link can be provided as required. The serial bit stream is buffered and converted to a parallel stream of information words. The buffer which is a part of each demodulator unit consists of three components:

- Serial input to parallel output (SIPO) register,  $b_n$  bits in length ( $b_n = (sL)/n$  where  $(sL)$  is the number of bits/burst and  $n$  is an integer).
- A  $b_n$  bit clock counter which transfers the contents of the SIPO register to a  $b_n$  bit wide parallel holding register at the end of each clock interval.

Clocking rates range from  $T_1$  to  $K \times T_2$ , as previously mentioned. If  $K = 10$  is assumed as an example, clocking would occur at 63.12 Mbps ( $10 \times 6.312$  Mb/s). Shift registers and counters capable of operating at this speed are readily available using Schottky or emitter-coupled logic (ECL) families. Because of the lower clocking/transfer rate, the  $b_n$  bit holding register can be made from regular transistor-to-transistor logic (TTL) devices.

The output from all demodulator buses is multiplexed into a single stream of information words and passed to the digital switch interconnect subsystem.

#### 4.2.2 Digital Switch Interconnection

The switch interconnection subsystem, which is the core of the processor, performs the function of interconnecting the various elements of the processing system. The switch is assumed to operate in a time division mode with interconnection between devices determined by the switch supervisor. Address information is obtained from the demodulation processor module and from the switching algorithm provided by the ground control station.

Memories 1 to R are used to organize the uplink data according to their destination. While data are in the memory, further operation can be performed on them by the word processors attached to the switch. Functions such as decoding, re-encoding, ARQ strategies, and one to many type transmission can be handled by the word processors.

The interconnection subsystem is a non-blocking switch. The term non-blocking is usually omitted since one characteristic of the switches used in multiprocessor systems is that they are complete with respect to the memory units (i.e., there is a separate bus associated with each memory, and the maximum number of transfers that can take place simultaneously is limited by the number of memory boxes and the bandwidth-speed product of the buses, rather than by the number of paths available).

The important characteristics of a system utilizing an interconnection matrix are the extreme simplicity of the switch-to-functional unit interfaces and the ability to support simultaneous transfers for all memory units. To provide these features requires major hardware capabilities in the switch. Not only must each interconnect point be capable of switching parallel transmissions, but it must also be capable of resolving multiple requests for access to the same memory module occurring during a single memory cycle. These conflicting requests are usually handled on a predetermined priority basis, e.g., input/output has highest priority in all conflicts, processor number 2 has primary access priority to memory 2, etc. The result of the inclusion of such a capability is that the hardware required to implement the switch can become quite complex.

Microprocessors attached to the switch handle the data transfer between demodulators and modulators and the memory which is partitioned for each link. The connectivity algorithm used to drive these microprocessors is obtained from a central processor which receives assignment information from the ground control station via the command and control link.

Processing speeds for T1 and T2 scanned input (uplink) data handling are considered to be low to moderate. Processing rates are reduced by serial-to-parallel (S/P) conversion circuitry, since it would be prohibitive to handle all uplink data bits on a bit-by-bit basis. Multi-bit word transfers to memory can be accomplished in 50 ns or less using currently available bipolar random access memory technology.

#### 4.2.3 Downlink Modulators

Data destined for a specific downlink are passed from memory to the respective modulator through the switch as a series of information words. A parallel to serial (P/S) shift register converts the words in serial bit stream. Data are passed along a one bit bus to the modulator. Here, as with the uplink, one bit processors are available to perform specific functions required by the link.

TDM downlink processing speeds are high. In a typical case, the translation from N uplink beams, each with M FDMA channels to N downlink beams using TDM, causes the downlink channel rate to be at least M times the uplink rate. If a downlink is scanned, the rate goes up by the ratio of scan frame time to user dwell time, K. For values of 10 for M and 100 for K, the downlink rate for T2 carrier approaches 1000 times the uplink rate (~6 gigabits/sec) which may be feasible in the future [Bosch, 1979].

P/S circuitry aids assembly of data words unloaded from the user address in the common memory for downlink transmission. Word transfer takes place at a very high rate, which can only be accomplished by the use of wide words in the memory organization of the processor. Wide-word memory organization reduces output processing data transfer rates to small values.

#### 4.2.4 Word Processor Architecture

Controller processor architecture can be implemented with current bit-slice bipolar devices similar to the INTEL 3000 or AMD

2900 series. Large instruction and data bus word widths can be assembled with these devices. Utilization of interleaved and/or pipelining techniques allows word processing rates of 20 megawords/sec. If gigabit technology [Bosch, 1979] can implement logic gate arrays similar to those in present bipolar slice devices, greater than 1 gigaword/sec rates may become available.

Word processors will probably use bipolar designs but 16 bit microprocessor devices (such as the INTEL 8086, Motorola 68000, or Zilog 8000, soon to be readily available) are possible alternative candidates. Another alternative processor structure uses direct-memory-access (DMA) techniques involving high-speed transfer of large blocks of accumulated data (multiple sub-slots) from each type of user. DMA controllers, operating with an input and/or output controller peripheral, pass moderate speed information from input (uplink) to output (downlink) across a common (wide) data bus. Instead of using a bulk memory design, memory would reside in the random access memory (RAM) serviced by DMA and peripheral devices. A device such as an AM 9517 (made by AMD, Sunnyvale, CA) can perform this function at 2 megawords/sec.

#### 4.3 MEMORY REQUIREMENTS AND TECHNOLOGY

A large high-speed memory (estimate at around 100M bits) is required in the processor section. The quantity is dependent on the total uplink capacity received per master system timing frame and the amount of delay is incurred in the satellite before beam connectivity is established.

As an example, assume a beam, composed of 10 FDMA channels each with T1 data, scanning 100 uplink beam coverage areas during a frame of 10 msec; storage required for a whole frame is at least 15.44M bits (100 areas x 10 channels/beam x 1.544 Mb/s (T1) x 0.01 sec). For a double buffer arrangement, the bulk memory would be twice this value or 30.88M bits. This amount of memory is moderate but the 15.44M bits must be transferred in 0.01 seconds which is a 1.544 Gb/s link transfer rate.

For one T1 channel per user, 15,440 (1.544 Mb/s x 0.01 seconds) user bits (not including overhead) are received during each master frame. The basic buffer portion of the bulk memory for this user would probably have to be at least twice the calculated size because of processing delay. The uplink is received continuously and must store a frame of data before sending it down as a high-speed TDM burst. A basic user memory storage allocation is therefore 30,880 bits. The bulk memory would then be organized as 1000 blocks of



memory (one for each uplink user) each containing at least 30,880 bits. For T2 users, 4000 such locations would be needed. Therefore, T2 operation requires a total memory of 123.52M bits (organized as 4000 blocks, each containing at least 30,880 bits). Additional information such as user address tables and input/output access address counters/pointers would probably raise each section requirement to 32K bits. A hardware implementation could use 4000 four kilobyte RAM integrated circuits. For an assumed 0.1 W power dissipation per package and including 2000 integrated circuits needed for support, the total power dissipation of the processor would be at least 0.6 kW.

Various technologies, such as core, disk, tape, and plated wire, are available to handle bulk data requirements; however, all suffer from various undesirable attributes such as slow access time, moderate support circuitry requirements, and moderate-to-high power dissipation. Figure 4-4, taken from [Theis, 1977] shows access speed vs. memory capacity for various memory technologies.

RAM memory devices using bipolar, metal oxide semiconductor (MOS), magnetic bubble domain, or charge-coupled device (CCD) technology can meet memory capacity requirements and in some cases moderate-to-fast access time requirements. Power dissipation is a major factor. Bubble memories and CCDs provide relatively low cost mass storage in a serial format, but suffer from access time delays due to serial register storage lengths (typically 256 bits) and low clocking rates (<500 kHz). Currently, individual devices are able to store 256K bits and multi-megabit boards are being produced. A typical device being produced by Rockwell International (Anaheim, CA) is identified as a type BBM256 and contains 260 bubble loops of 1025 bits each. Clocking rates are 150 kHz maximum. Each device dissipates 820 mW. A 1 megabyte memory would consist of 32 packages and dissipate 25 W exclusive of support circuitry. CCD components have similar characteristics. Assembly of many of these devices into a very large (<100M bit) memory is certainly feasible but access time and power dissipation would be a problem. Redesign of these devices as FIFO buffers to handle faster clocking rates (>10 MHz) could provide large word width buffers for individual user memory allocations. N-type MOS (NMOS) dynamic memories with 64K bits per package are also available but with access times of 500 ns.

Bipolar RAMs can provide fast access times. Current technology is capable of packaging 4K bits of memory in a package with 85 ns access times (Intel 2147). RAMs under development by IBM, Essex Junction, VT [IBM, 1979] utilize fault-tolerant designs and E-beam fabrication techniques for 8-bit wide RAMs in 18, 32, and 64K-bit



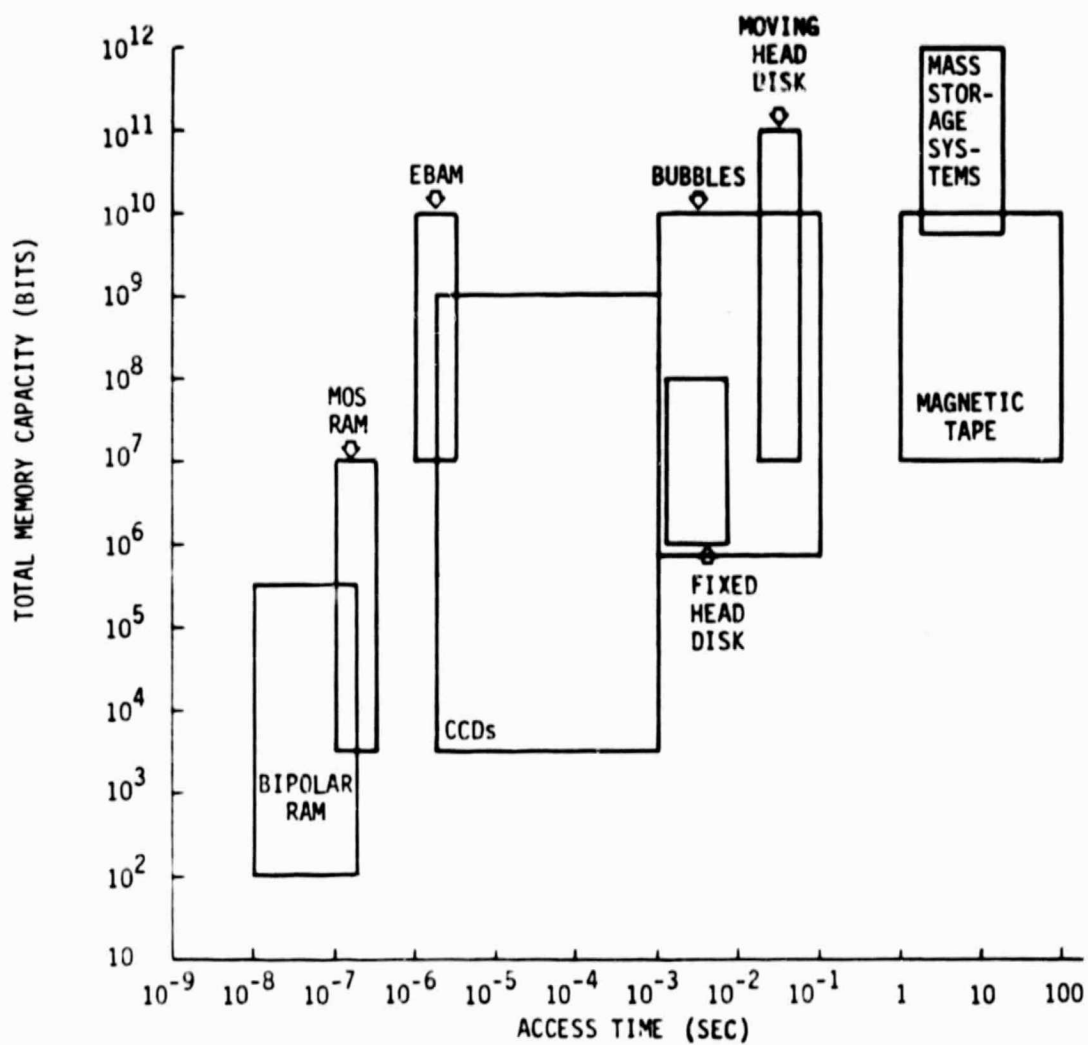


Figure 4-4. Memory Capacity Versus Access Time

sizes. The 64K device is externally organized as 8K x 8 and can operate at a 1.25 MHz byte output rate.

Maximum obtainable density in memories is directly related to the memory circuit (one, two, or three transistors per bit) and power dissipation ratings of the package. Very high density (64K) NMOS memories which require periodic refreshing of the data every one or two milliseconds have been designed using one transistor per bit. Static memories (not needing refreshing) are usually less densely packed (two or three transistors per bit) and are available in a 4K configuration. Current packaging designs are limited in overall power dissipation capabilities. Typically, common integrated circuit packages are rated at 1 W dissipation or less. Larger package designs, containing multiple chips or a larger high-density die, are being investigated by several manufacturers. Gigabit memories suitable for spacecraft application are described in a brief article presented at the 1978 International Telemetry Conference [McDermott, 1978]. Relatively large memories which have already been flown in space are presented in table 4-1.

#### 4.4 LOGIC SPEED REQUIREMENTS

Integrated circuit logic devices are made from arrays of basic logic gates such as ANDs, OR, NORs, etc. The operating speed of an array is dependent on the functional propagation delays (input-to-output) developed in each gate. A family of curves showing the maximum permissible gate delay versus required input clocking rates for multiple gates in series is presented in figure 4-5.

Commercially available logic gates, flip-flops, and shift register devices that can be clocked at 600 MHz rates are available from Plessey Electronics in England and Motorola in the United States. These devices employ ECL technology. Development devices operate at multi-gigabit rates.

The power-speed product is an important figure of merit for logic devices or arrays made from the various gate technologies (TTL, ECL, NMOS SOS,  $1^2$ L, GaAs FET). The power-speed product is usually expressed in picojoules and is formulated by the product of the power dissipation of the simplest gate function developed for the technology involved and the propagation delay from input to output through the gate. Speed power products as a function of gate delay for various current logic families are presented in figure 4-6. Logic devices will have to be manufactured as arrays to minimize capacitance effects and reduce switching speeds below 1 ns

Table 4-1  
On-Board Space Computer Systems\*

EVOLUTION OF COMPUTING SPEED AND INTERNAL MEMORY CAPACITY			
YEAR	COMPUTER APPLICATION	COMPUTING SPEED (kb/s)	CAPACITY (WORDS)
1960	SATURN I	3.0	3,644 (24-bits)
1963	GEMINI	7.0	4,096 (39-bits)
1964	SATURN IB and V	11.3	16,384 (26-bits)
1971	SKYLAB	60	16,384 (16-bits)
1974	SPACE SHUTTLE	325 <sup>a</sup> 480 <sup>b</sup>	40,960 (32-bits)
1976	USAF <sup>c</sup>	200	61,000 (32-bits)
(a)	Floating Point		
(b)	Fixed Point		
(c)	Raytheon Company		

\* A. E. Cooper, W. T. Chow, "Development of On-Board Space Computer System," IBM J. Res Develop., January 1976.

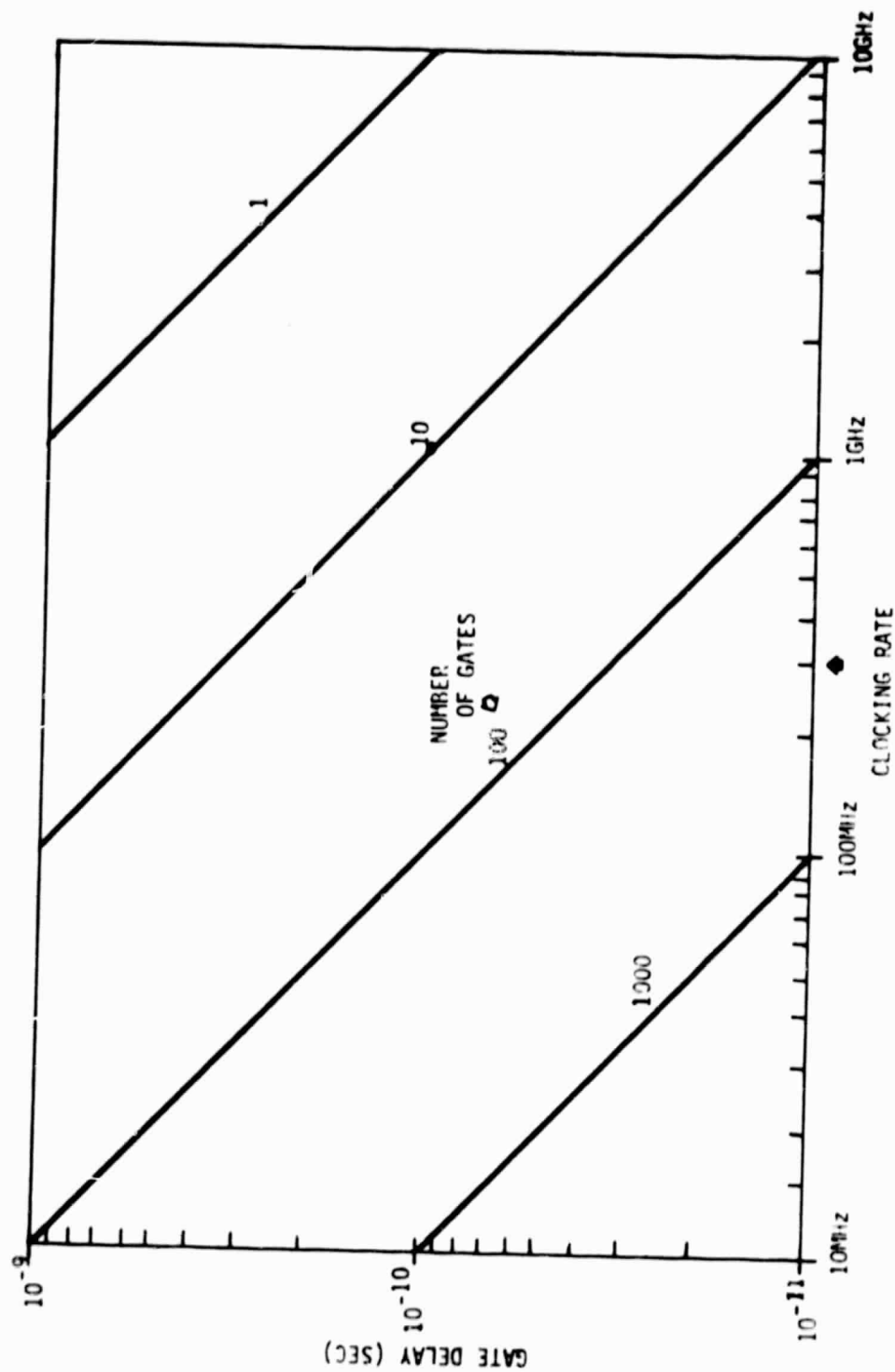


Figure 4-5. Maximum Permissible Gate Delay as a Function of Clocking Rate

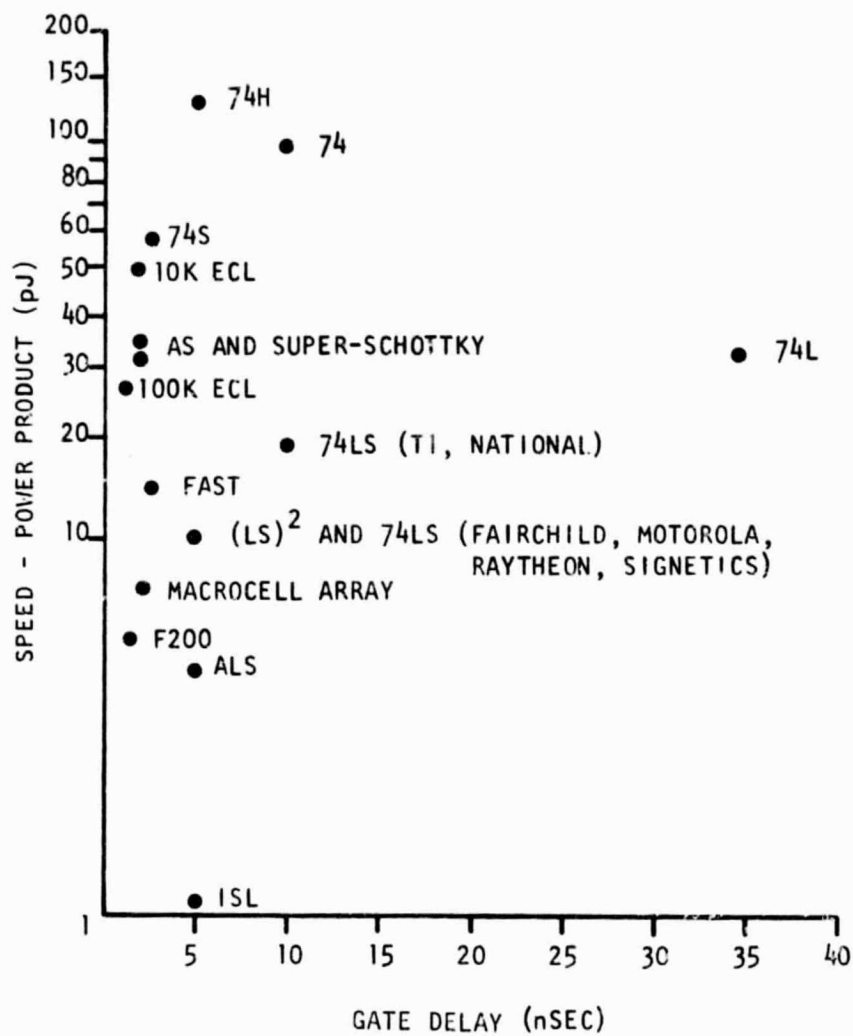


Figure 4-6. Speed-Power Product for Various Logic Families [Ranaaga, 1979]

(1000 ps). Future array characteristics are presented in figure 4-7. As the number of gates in an array increases, power dissipation of the array (chip) will become especially important.

An excellent bibliography and review of what high speed digital electronics is and what will be available is given by [Bosch, 1979]. The article describes integrated circuit technology capable of handling data rates of several Gbps.

High-speed multipliers are an example of the application of high speed logic to data processing. Multipliers require continual shifting and adding of data at fast rates. In [Bosch, 1979] references are made to multipliers operating at 40 MHz (8 x 8 bit) using Josephson Junction technology and at 100 MHz (16 x 16 bit) using GaAs field effect transistor (FET) devices, and to multi-bit adders with carry propagation delays of 50 to 140 ps. Characteristics of several multiplier devices currently available today are presented in tables 4-2 and 4-3. Tables 4-4 and 4-5 present high speed arithmetic devices. Continued development will impact demodulation/remodulation, input/output, memory and processor devices used in advanced processing satellite systems.

#### 4.5 SUMMARY

- A digital processor capable of performing high speed switching of multiple T1 and T2 uplink/downlink channels appears to be feasible in the 1990 to 2000 time frame.
- Gigabit/sec digital device technology under development may be useful in large memory arrays.
- Although multi-processor system designs are being developed, the architectural configuration for efficient multiport memory with a large number of ports and digital switch interconnecting subsystems needs to be further developed.
- Digitally programmable multiple high rate all-digital demodulator/modulator modules need to be developed.
- Designs which reduce power requirements and which can easily be checked for proper operation using built-in self test algorithms should be developed.
- Fault-tolerant design techniques should be applied in each area so as to maximize on-orbit lifetime and usefulness.

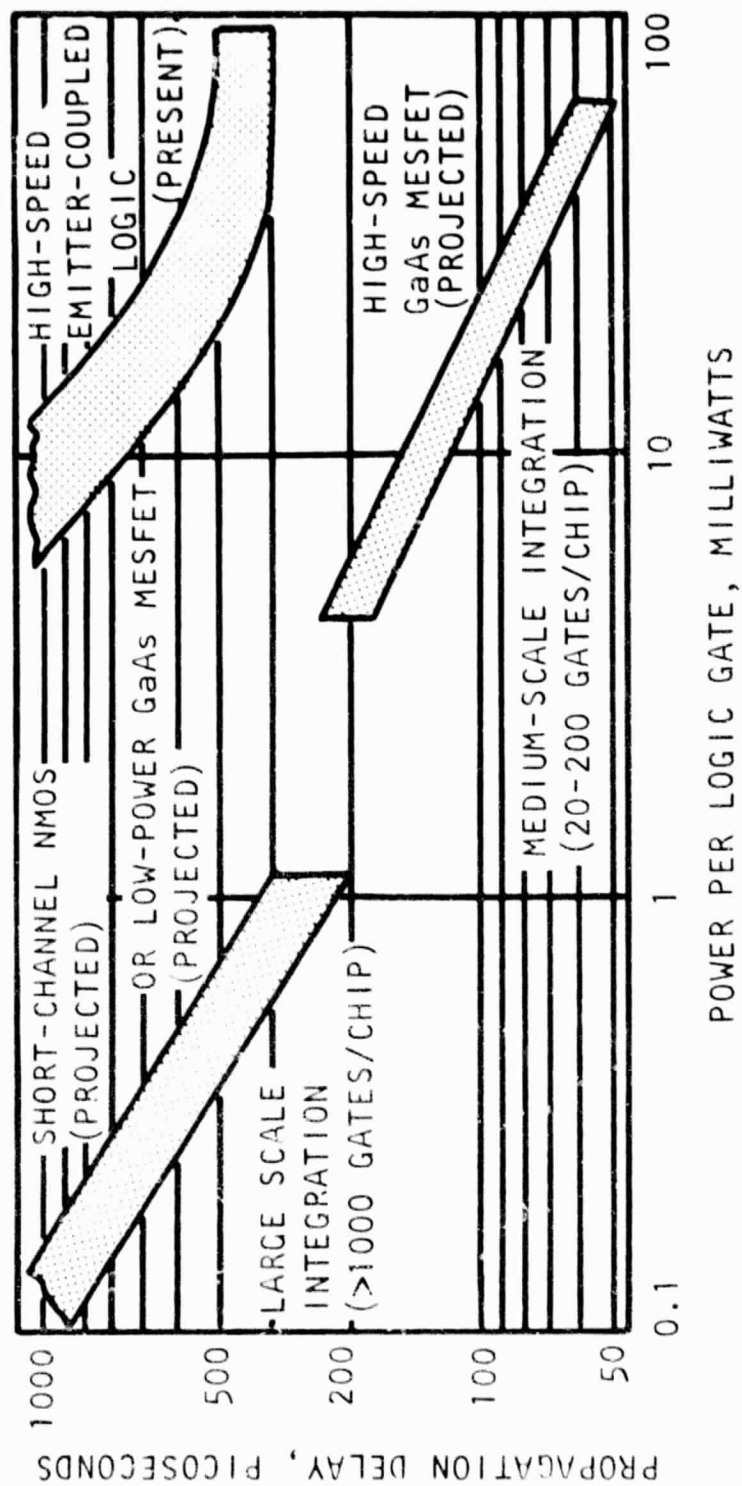


Figure 4-7. Present and Projected Status of Subnanosecond Logic

### Comparison of Common Technologies Used in Implementing High Speed Arithmetic Devices

TECHNOLOGY/ YEAR		GATE CHARACTERISTICS					2-INPUT EXCLUSIVE-OR		
INTRODUCED	FUNCTION	DELAY	POWER	SPEED-POWER PRODUCT	DENSITY (gates/mm <sup>2</sup> )	DELAY	POWER	SPEED-POWER PRODUCT	COMMENTS
ECL-1111 (1968)	NOR	1.1 ns	60 mW	66 pJ	30	1.3 ns	70 mW	91 pJ	Limited number of functions
ECL-1000 (1971)	NOR	2 ns	25 mW	50 pJ	30	2.5 ns	50 mW	125 pJ	Large selection of functions
S/TTL (1970)	NAND	3 ns	20 mW	60 pJ	30	7 ns	60 mW	420 pJ	Large selection of functions
LS/TTL (1972)	NAND	10 ns	2 mW	20 pJ	30	10 ns	8 mW	80 pJ	Large selection of functions
IIL (1975)	NAND	10 ns	0.1 mW	1 pJ	300				Not a mature technology
NMOS (1973)		100 ns	0.1 mW	10 pJ	130				For reference only
EEIC (1977)		0.25 ns	2 mW	0.5 pJ	100				Still in research and development



Table 4-3

## Comparison of Available Multiplier ICs

VENDOR/ DEVICE	CONFIGURATION	PINS	SPEED	POWER	DATA CODE	AMOUNT OF PARALLELISM	ALGORITHM
TRW/ MPV-8	8 x 8	40	130 ns	1.2 W	2's comp	Full	AND gates and carry-save adders
TRW/ MPV-12	12 x 12	64	150 ns	3.5 W	2's comp	Full	AND gates and carry-save adders
TRW/ MPV-16	16 x 16	64	180 ns	5 W	2's comp	Full	AND gates and carry-save adders
MMI/ 6755R	8 x 8	40	100 ns	1 W	2's comp/ unsigned	Full	Modified Booth's; modified Wallace Tree
MMI/ 67516	16 x 16	24	800 ns	1 W	2's comp	Multiplicand/ 2 multiplier bits	Modified Booth's
MMI 67508	8 x 8	20	400 ns	0.75 W	2's comp	Multiplicand/ 2 multiplier bits	Modified Booth's
AMD/ 25S05	2 x 4	24	25 ns	0.6 W	2's comp	Full	Booth's
AMD/ 25LS14	8 x 1	16	50 ns	0.5 W	2's comp	Multiplicand/ 1 multiplier bit	Booth's
AMD/ 25LS2516	8 x 8	40	400 ns	1 W	2's comp	Multiplicand/ 2 multiplier bits	Modified Booth's
TI/ 74S274 (ROM)*	4 x 4	20	50 ns	0.5 W	Unsigned	Full	ROM lookup table*
Motorola/ 10183	2 x 4	24	20 ns	0.8 W	2's comp	Full	Booth's

\* When used with companion device 74S275, partial products are added in a Wallace Tree configuration

Table 4-4  
Comparison of Arithmetic Logic Units

<u>PART NO.</u>	<u>FUNCTION</u>	<u>GATE COUNT</u>	<u>SPEED</u>	<u>POWER</u>
74S181	4-bit ALU	75	11 ns	600 nW
10181	4-bit ALU	75	7 ns	600 mW
74S281	4-bit ALU/ accumulator	100	22 ns	700 mW
74S182	4 groups carry-lookahead	20	7 ns	350 mW
10179	4 groups carry-lookahead	20	4 ns	300 mW

COMPARISON OF SPEED/POWER AT SYSTEM LEVEL

<u>PART NO.</u>	<u>TECHNOLOGY</u>	<u>4 BITS</u>	<u>SPEED/POWER</u> <u>16 BITS</u>	<u>64 BITS</u>
74S181/ 74S182	S/TTL	11 ns/600 mW	18 ns/2.7 W	28 ns/11 W
10181/ 10179	ECL	7 ns/600 mW	11 ns/2.7 W	17 ns/11 W

Table 4-5

Performance Comparison Between 8 x 8 and 16 x 16 Multiplication

VENDOR/ NUMBER	CONFIGURATION	PINS	8 x 8 MULTIPLICATION			16 x 16 MULTIPLICATION		
			NO. OF PACKAGES	SPEED	POWER	NO. OF PACKAGES	SPEED	POWER
MMI/ 67558	8 x 8	40	1	100 ns	1 W	14*	140 ns	9 W
TRW/ MPY-8	8 x 8	40	1	130 ns	1.8 W	14*	170 ns	10 W
Motorola/ 10183	2 x 4	24	8	50 ns	6.4 W	32	100 ns	25.6 W
AMD/ 25S05	2 x 4	24	8	75 ns	5 W	32	150 ns	20 W
TI/ 74S274	4 x 4	20	12**	75 ns	5.4 W	45	120 ns	21 W
TRW/ MPY-16	16 x 16	64	-	-	-	1	180 ns	5 W
MMI/ 67516	16 x 16	24	-	-	-	1	800 ns	1 W
AMD/ 25LS2516	8 x 8	40	1	400 ns	1 W	2	800 ns	2 W

\* 4 packages are 8 x 8 multipliers, 10 are adders (74S181/10181)

\*\* 4 packages are 4 x 4 multipliers, 8 more are Wallace Tree bit-slices (74S275)

#### REFERENCES

Bosch, B. G., "Gigabit Electronics - a Review," Proc. of the IEEE, Vol. 67, No. 3, March 1979, pp. 340-379.

Elson, B. M., "Computer Development Time Being Cut," Aviation Week & Space Technology, 26 February 1979, pp.67-70.

Enslow, P. H., Jr., "Multiprocessor Organization - A Survey," Computing Surveys, Vol. 9, No. 1, March 1977.

IBM, Electronic Design, No. 6, 15 March 1979, pp. 32-33.

McDermott, Electronic Design, No. 24, 22 November 1978, pp. 25-26.

Ranada, David, "High-Performance Digital Logic," Electronic Design News, 5 February 1979, p. 84.

Theis, D. J., "Memory Technology Update," Report No. TOR-0078 (3475-20)-2, The Aerospace Corporation, 28 December 1977.

## SECTION 5

### MICROWAVE SWITCH TECHNOLOGY

Advanced signal processing concepts for future satellite communications systems can utilize RF processing architecture. This section is devoted to evaluating and making recommendations concerning the RF switch which is viewed as the heart of an RF processing satellite communications system.

#### 5.1 RF SWITCHING ARCHITECTURE

The switching of large numbers of microwave channels is a new and specialized subject. A number of techniques can be marshalled to support the design of such a switch. Some of these techniques are quite mature and are the subject of individual textbooks, while some are new and the object of intensive research. The scope of the present study was chosen to answer the following questions:

1. What technical obstacles impede the development of a large capacity RF switch for use in the latter half of the 1990's?
2. How should NASA/industry proceed to ensure that such a switch will be realized?

In carrying out the study, conferences were held with many of the key experts in the United States. (A list of these conferences is presented in figure 5-1.) The literature was searched for information on switching and supporting techniques.

The requirements for an RF switch are determined by the communications environment described in section 2.5. It is postulated that the bandwidth in any fixed beam will be 2.5 GHz. This will probably be broken up into 500 MHz bands to accommodate transponder technology. Thus, there may be up to five ground station concentrators in some of the metropolitan areas served. In consequence, the number of input ports (and therefore output ports) on the switch may well exceed the number of fixed beam cities plus

HARRY DIAMOND LABORATORIES	(R. Garver)
OMNI SPECTRA CORPORATION	(B. Robinson)
A.F.A.L. WRIGHT PATTERSON FIELD	(E. Nichols, J. Adair)
FORD AEROSPACE	(R. Davies, R. Jorash)
HUGHES AIRCRAFT	(B. Dobratz, R. Sakamoto)
TRW	(H. Poza, W. Nations)
MICROWAVE ASSOCIATES	(D. Fryklund)
DEXCEL CORPORATION	(P. Ho)
ALPHA CORPORATION	(K. Kelley)*
R.C.A. ASTRO	(F. Drago)*
COMSAT CORPORATION	(F. Assal)
RADC HANSCOM AFB	(A. Yang)
ELECTROMAGNETIC SCIENCES	(J. Pippin)*

---

\* PHONE CONFERENCE

Figure 5-1. Organizations Contacted

the number of phased array pairs, since different bands in the same beam may be destined for separate cities as indicated in figure 5-2.\* Attention is therefore focused on a 100 x 100 (input port, output port) switch.

When microwave switches are small, where small may mean 5 x 5, they might be realized by extrapolation of techniques used in constructing switches commercially available. However, when they become large, new difficulties arise which call for designing the switch as a new entity at the outset. For example, a large switch may generate electromagnetic interference (EMI) among its own channels, control line layout may physically interfere with information channels, cooling may be more difficult to route, etc. It is therefore necessary to study a large switch so that these problems may be brought to light, and techniques, or at least approaches, may be devised for their solution. A 100 x 100 switch is large enough to provide assurance that the problems will be encountered even though the exact boundary between large and small may be indistinct.

All of the 500 MHz channels in a given beam may have a common destination. Because flexibility in connecting any 500 MHz channel input to any output is required, all channels in the switch and their interconnections must be able to pass the full 2.5 GHz bandwidth.

The first two specifications for the microwave switch indicated are:

1. Capacity: 100 x 100 minimum
2. Bandwidth: 2.5 GHz (channels and interconnections)

The placement of a switch between the uplink and downlink portions of a transponder allows the switch to be designed to operate at some intermediate microwave frequency which is advantageous for switch realization. The technologies investigated suggest that this frequency should lie somewhere between 2 and 4 GHz

---

\* A study of the Rand McNally map of standard metropolitan areas shows there are 73 areas with populations in excess of 500,000. These could comprise the fixed beam cities.

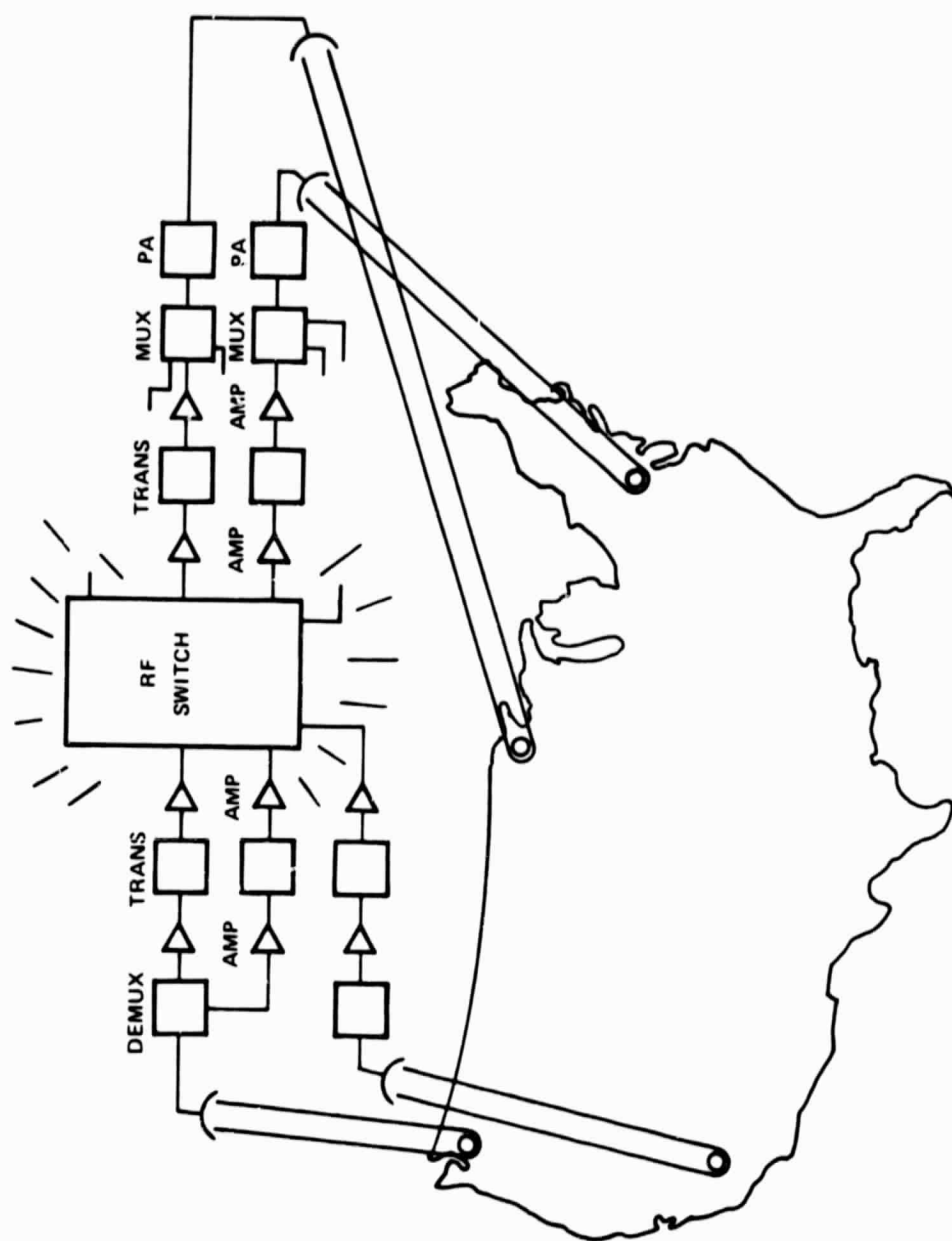


Figure 5-2. Interconnection of Fixed Geographical Regions



since circuit and device technologies perform well there. This will require downward translation to IF at say 3 GHz, switching, and upward translation to the downlink frequency.

The RF switch is assumed to operate in conjunction with a TDMA system. The set of interconnections existing at a given instant of time is called a switch state. The switch undergoes a succession of states over a period of time called a frame. When the frame is completed, the switch repeats the same succession of states periodically. It consequently executes frame after identical frame for the purpose of supporting all the channels in the country communicating through the satellite. The organization of frames is illustrated in figure 5-3.

If all slots were the same size, figure 5-3 would look different. All of the interslot boundaries would line up one below the other and the 100 channels would change together at each slot. This would require a larger peak load of the switching power supply. More important, the slot size would have to be frozen at a duration in microseconds required by the highest capacity traffic link. A link with much less traffic would get exactly the same slot size and, if the system had a uniform transmission rate, the link source would complete its transmission well before the slot interval had elapsed. For the rest of the slot, nothing would be transmitted and the link would be maintained while not being used. This represents a loss of throughput and consequently of revenue.

The unequal slot durations enable the system to be flexible in allocating resources based on traffic demand. The slot boundaries therefore do not coincide and few switches are thrown at the same instant. This has the advantage of smoothing the driver power flow necessary for switch operation. The flexibility offered by variable slot size has the advantage of utilizing satellite communications capability in the most efficient way by matching it exactly to the load distribution.

In an examination of the character of the TDMA frame and slot operation, certain additional switch requirements became apparent. These are:

1. Switching time 50 ns maximum
2. Slot size approximately 2  $\mu$ s minimum
3. Slots of unequal size
4. Non-simultaneous operation.

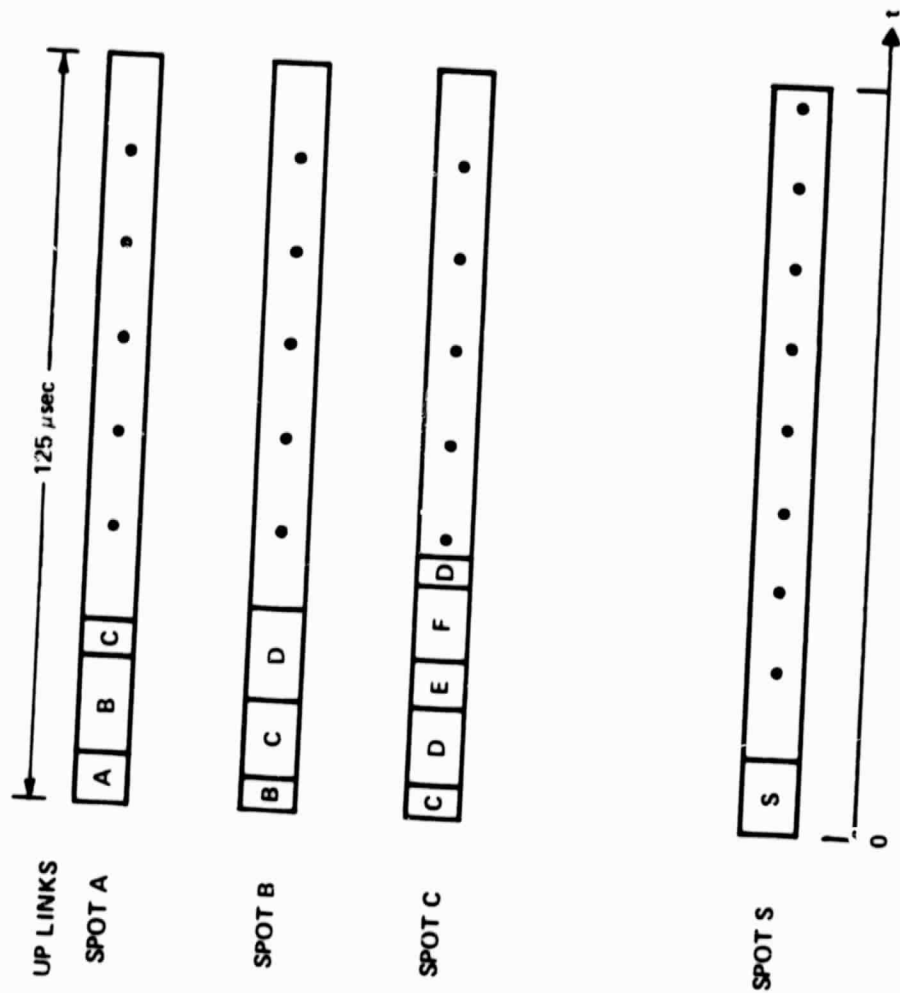


Figure 5-3. Frame Slot Organization

#### 5.1.1 Switch Technologies

An RF switch will consist of combinations of five elements. These are:

1. Transmission Lines
2. Devices (switching)
  - a. time varying (such as a switch)
  - b. active (such as an amplifier)
3. Drivers
4. Logic
5. Structure

Logic is not considered a critical design problem; current design practice is capable of meeting the requirements that will be set in this area. A large number of drivers have been designed for phase shifters and simple switches, so there is art available. There is a potential problem in power consumption by switch drivers as shown below.

The principal areas of concern are forms of transmission media, and switching devices for microwaves. There has been an evolution in microwave transmission media which has led to widespread use of various microwave printed circuits. Figure 5-4 shows three forms of these plus an optical medium which is discussed separately. Printed circuits are favored for microwaves because of their compact form, light weight, and ease and economy in fabrication.

The oldest of these media, and still popular, is stripline. This consists of a flat metallic center conductor supported between two ground planes. The medium between the center conductors and the two ground planes is a dielectric, as shown in figure 5-4, which most often serves to support the center conductor. The electromagnetic field has a distribution like that on a coaxial line in that the two ground planes are at the same potential so that field lines all converge on the central conductor or leave it. The characteristic impedance of a stripline depends on the ratio of the width of the center conductor,  $w$ , to the thickness of the dielectric between it and the ground plane,  $d$ .

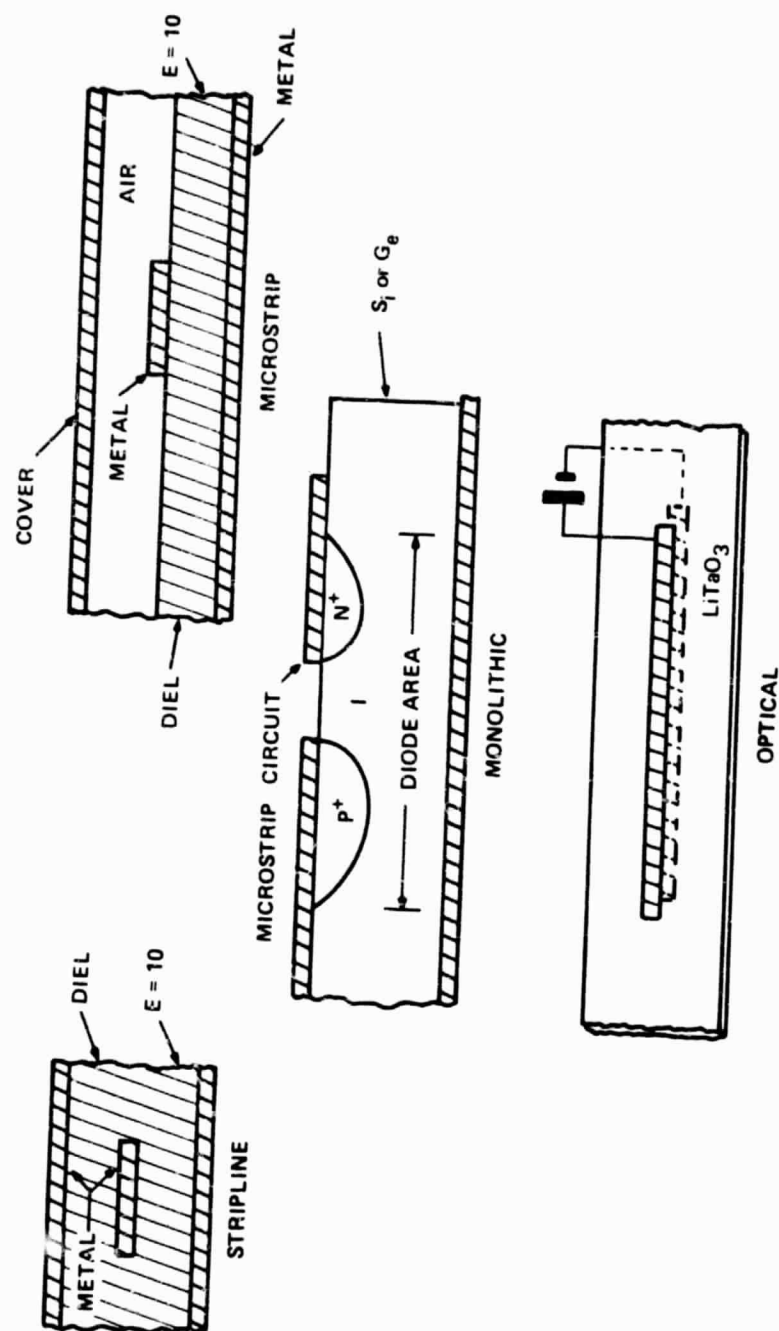


Figure 5-4. Candidate Media

The second and perhaps most widely used medium is microstrip in which the dielectric above the center conductor and the upper ground plane are dispensed with. The field is concentrated in the dielectric below the center conductor although there is some fringing in the space above. A cover is generally placed over the circuit to protect it, which has some slight effect on the characteristic impedance. Microstrip is preferred because of the accessibility of the center conductor and consequent ease of connecting devices to it.

The switch microstrip circuit may become monolithic in the near future, as shown in the central sketch of figure 5-4. The dielectric could then be composed of intrinsic semiconductor material having low loss. Active devices, such as the diode shown, or transistors could be diffused or epitaxially grown where needed. Connecting microstrip center conductors could then be deposited. The potential benefits of this approach would be the economy of integrated circuit production techniques and reduction in size of the switch. There are some problem areas which require study including loss in the intrinsic material, isolation, and assembly into a large switch.

A third form of transmission line which is not an RF line at all is shown at the bottom of figure 5-4. It assumes that the signal has been used to modulate a beam of laser light, and the light conductive material is a crystal of lithium tantalate. The drawing shows the formation of an optical waveguide in the crystal by the application of a static electric field between two electrodes. The difference in index of refraction caused provides the guiding medium.

Not shown in figure 5-4 is the medium of empty space. This could provide a basis for a switch if that switch were a phased array providing multiple up and downlink beams. Each of these media are discussed later in connection with the application of suitable switch device technologies. These technologies are:

1. Phased Arrays with Multiple Beams
2. Pin Diode Switching
3. Ferrite Switching
4. Optical Switching

## 5. Metal Epitaxial Semi-conductor Field Effect Transistor (MESFET) Switching

The technologies are discussed after some further ground work is provided for switch architecture.

### 5.1.2 Switch Architecture

Switch architecture determines specifications for the switch, affects modes of switch operation, helps to set reliability levels, and determines size, weight, parts count, and ease of fabrication. Important architectural principles are evident in the elementary cross bar requirement shown in figure 5-5, which illustrates a set of switch functions.

The intersection of input and output transmission lines is called a cross point. Interconnecting the input and output lines at a cross point is called closing the cross point or making the connection. Through its construction, the switch is subject to certain constraints with regard to the operation of its cross points. Many different constraints are discussed in recent review articles [Joel, 1977; Marcus, 1977]. A network consisting of traffic concentrators operating on a TDMA basis via a switch has the following requirements:

1. Connection time slots are not equal.
2. When connectivity has been established between two cities which are bursting information to each other, interruptions caused by interconnecting other pairs of cities are unacceptable.
3. It must be possible to connect any input to any unoccupied output regardless of what other connections may exist at the time in the switch.
4. Input and output signals must be able to flow undisturbed at any time through a cross point which is not closed.

In addition, cross points must have sufficient isolation so that cross-talk does not become a problem when cross points are opened. It can be shown that the switch isolation performance, or ratio of insertion loss closed to insertion loss open, in decibels, should be equal to the desired signal to cross-talk ratio plus the number of inputs in decibels.

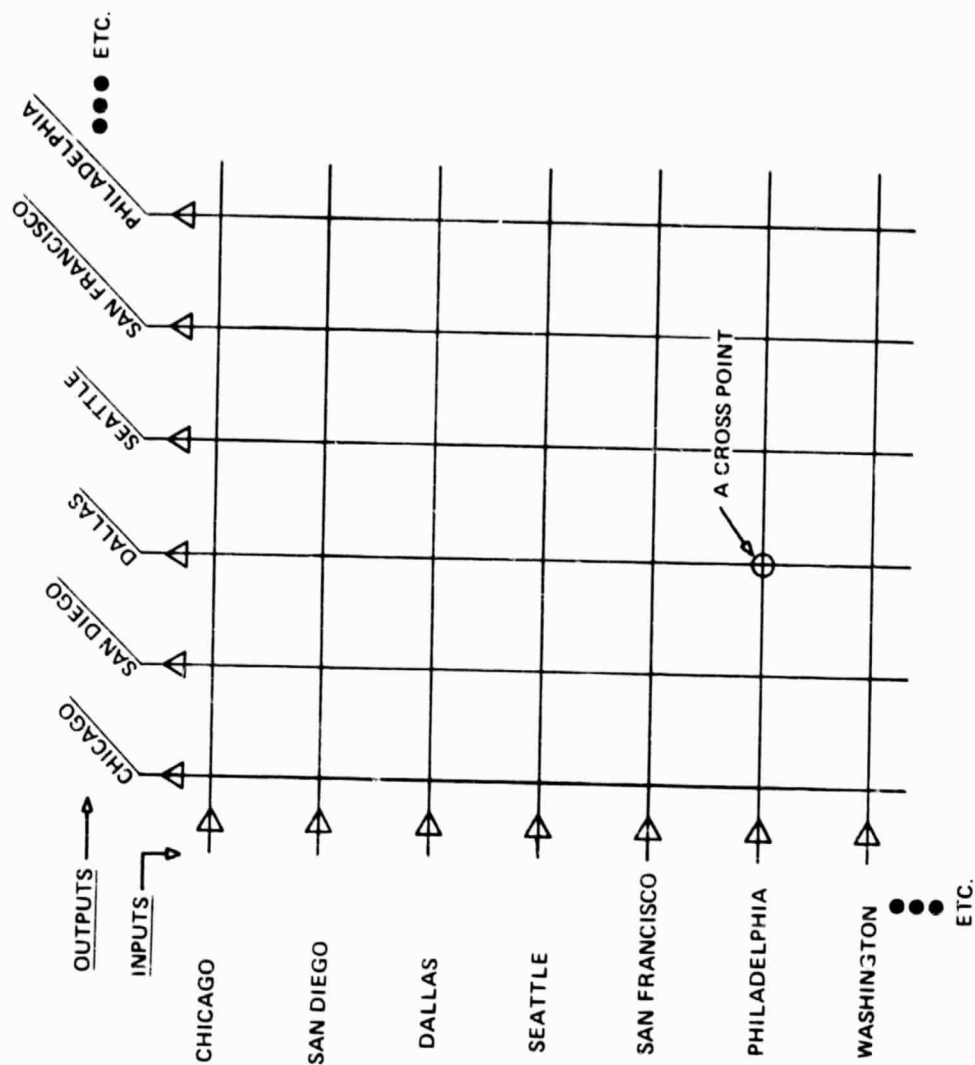


Figure 5-5. Cross Bar Requirement

It is highly desirable that the output signal levels from the switch be approximately equal, independent of the path taken. This eases the dynamic range burden on the transmit chains and reduces the need for specialized power conditioning as a function of switch output channel number. It has been shown that, within reason, the insertion loss through the switch does not matter since it is not located in the transponder where it can have a first order effect on G/T or EIRP. (This principle was tacitly accepted by COMSAT and INTELSAT in the development of a coupler fan-out switch.) However, the switch transfer function must be compatible with satellite specifications, which are estimated as:

Attenuation Variation (WB) = 1 dB max/2.5 GHz

Attenuation Variation (NB) = 0.5 dB max/10 MHz

Group Delay Variation = 1 ns/2.5 GHz

Table 5-1 lists tentative specifications for the switch.

### 5.1.3 Switch Architecture Realization

Realization means building a circuit of physical components and adjusting it to perform as desired. This section discusses methods of switch realization that depend on architecture. Later sections are devoted to discrete physical devices.

#### 5.1.3.1 Fan-Out/Fan-In Designs

The lower diagram in figure 5-6 shows the fan-out/fan-in technique and demonstrates the proliferation of layering or cross-overs of lines. The object is to be able to connect any one of four unprimed inputs, 1, 2, 3, 4 to any of the outputs 1', 2', 3', 4'. The diagram shows the signal flow paths but not the switches. The method used is to split each input into four parts, one for each output. This is the fan-out stage. After the split, switches can be inserted to permit or inhibit signal flow. There are many schemes for directing the split parts to the numbered outputs. The scheme used in the figure directs the first part to the opposite numbered output and subsequent parts to the non opposite numbered outputs in order. For example, the fractions from input 3 go to outputs 3', 1', 2', 4' in that order.



Table 5-1

## Tentative Switch Specifications

Size, n, (input/outputs)	100 x 100 minimum
Bandwidth (channels, connections)	2.5 GHz
Switching time	50 ns maximum
Mode of Operation	TDMA
Frame Duration	125 $\mu$ s
Slot Durations	proportional to traffic, unequal
Frame/Slot Schedules	reconfigurable
Slot Size	2 $\mu$ s minimum
Slot Communications Efficiency	85 percent
Switching Occurrence	independent times
Accessibility	to any unoccupied output
Interruption of Estab. Conn.	none due to switching
Signal Flow, at Unmade Cross Point	input, output signals bypass it
Input Power Level	1.0 mW per channel
Switch Isolation Coefficient, A (Closed/Open)	S/C + n (in dB)
Attenuation Var.(f) (WB)	$\pm 1.0$ dB/2.5 GHz maximum
Attenuation Var.(f) (NB)	$\pm 0.5$ dB/10 MHz maximum
Insertion Loss Absolute	50.0 dB maximum
Insertion Loss Variation	6 dB port to port maximum
Group Delay Var.(f) (WB)	$\pm 1$ ns/2.5 GHz
One to Many Facility	provided
Redundancy	sufficient to overcome failure rate
Prob. No Connect. Failure	99.99 percent, 7 years
Size	$\leq 4$ percent 4-ESS design goal*
Weight	$\leq 4$ percent 4-ESS design goal*
Power Consumption	$\leq 4$ percent 4-ESS design goal*
Cost	$\leq 4$ percent 4-ESS design goal*

\*The new Bell System 4-ESS consists of many racks of equipment, occupies in excess of 30,000 ft of floor space, uses 500 kW of power supplied from a standard 140 V source, has an estimated weight of more than 250,000 lb, and an estimated cost in excess of \$30,000,000.

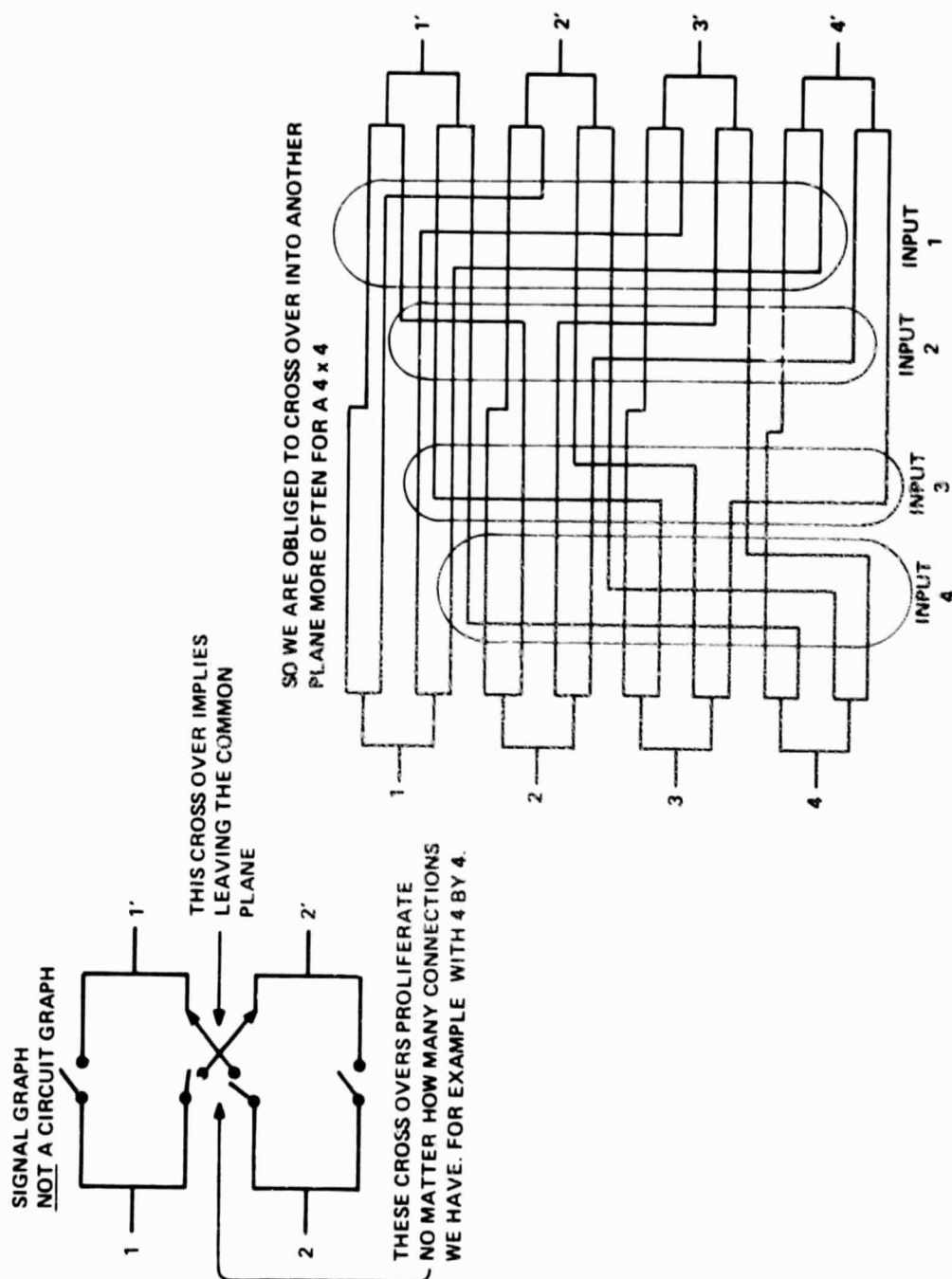


Figure 5-6. Fan-Out/Fan-In System

When the fractions arrive near outputs 1', 2', 3', 4', they may again encounter switches that determine which of the several available signals the output will accept. In any case, the arriving transmission lines must be fanned-in so that the signals can emerge from the destination port. As shown in figure 5-6, the fan-outs and fan-ins take the form of a corporate signal flow network. Thus, the objective of making a connection from every input to every output port is accomplished.

#### 5.1.3.2 Single Pole Multiple Throw Fan-Out Architecture

The diagram in figure 5-6 is not a physical representation. A better understanding of how such a switch is built up is gained from subsequent figures. Figure 5-7 shows a schematic of a single pole multiple throw switch (SPMT) which can be used as a basis for fan-out. A fan-out based on the single pole multiple throw switch is shown in figure 5-8 which expands a single input to 100 outputs. By means of switch selection, the entire input signal, except for attenuation in lines and through switching devices, is made to appear at only one selected output. This is equivalent to a single pole 100 throw switch.

The final physical realization of a 100 x 100 switch would use 200 of the units shown in figure 5-8. One hundred would be used as fan-outs and 100 as fan-ins. The entire assemblage appears as shown in figure 5-9. Each SP100T fan-out switch occupies a single horizontal layer with its input connector at the center of the rear edge. The 100 horizontal input fan-outs are stacked vertically in contrast to the 100 vertical fan-ins. The latter are positioned so as to be able to accept one output from a fan-out and thus to be able to connect their single output to any input. The 100 vertical fan-ins are stacked horizontally and a connection can be made from any input to any output.

Figure 5-10 is a further illustration of one physical arrangement for building a switch of the type described. A portion of a horizontal fan-out and a vertical fan-in are shown juxtaposed so that the right hand output of the fan-out is intended to enter the top input of the fan-in. A connector, not shown, would be used to bridge between the two planar circuits. Each circuit consists of microstrip transmission lines printed on a dielectric supported by a metal ground plane. To simplify the drawing, the fan-out and fan-in are shown as SP8T switches.

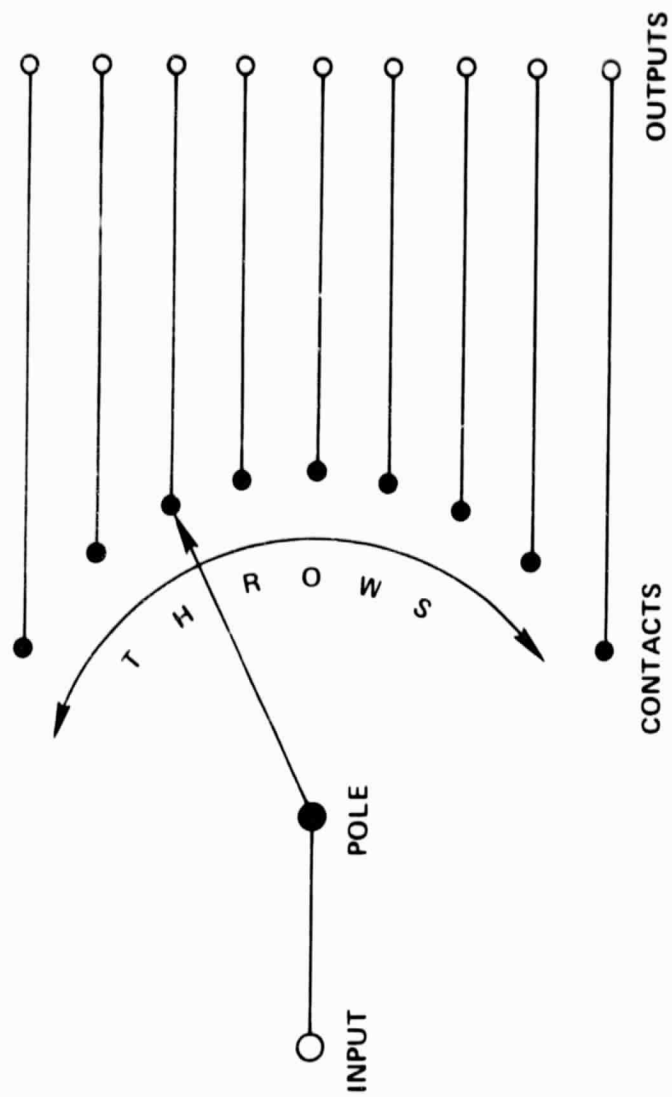


Figure 5-7. Single Pole Multiple Throw Switch

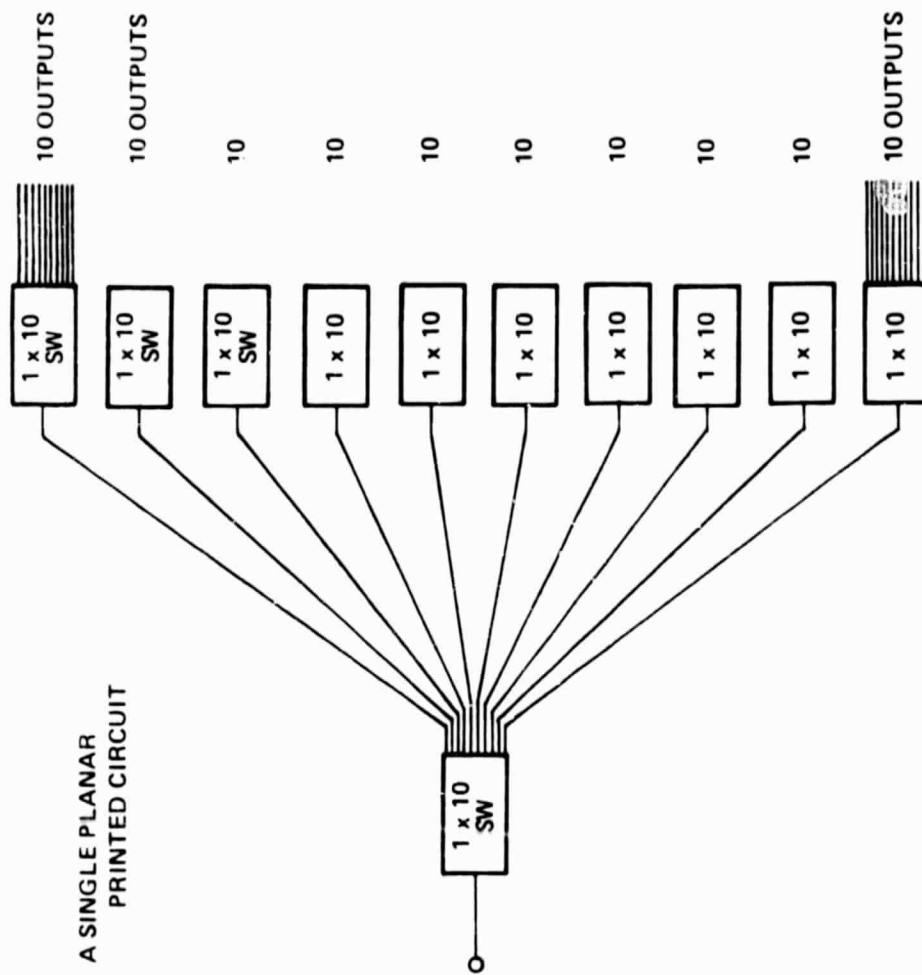


Figure 5-8. A 1 x 100 Switch Made of 1 x 10 Switches

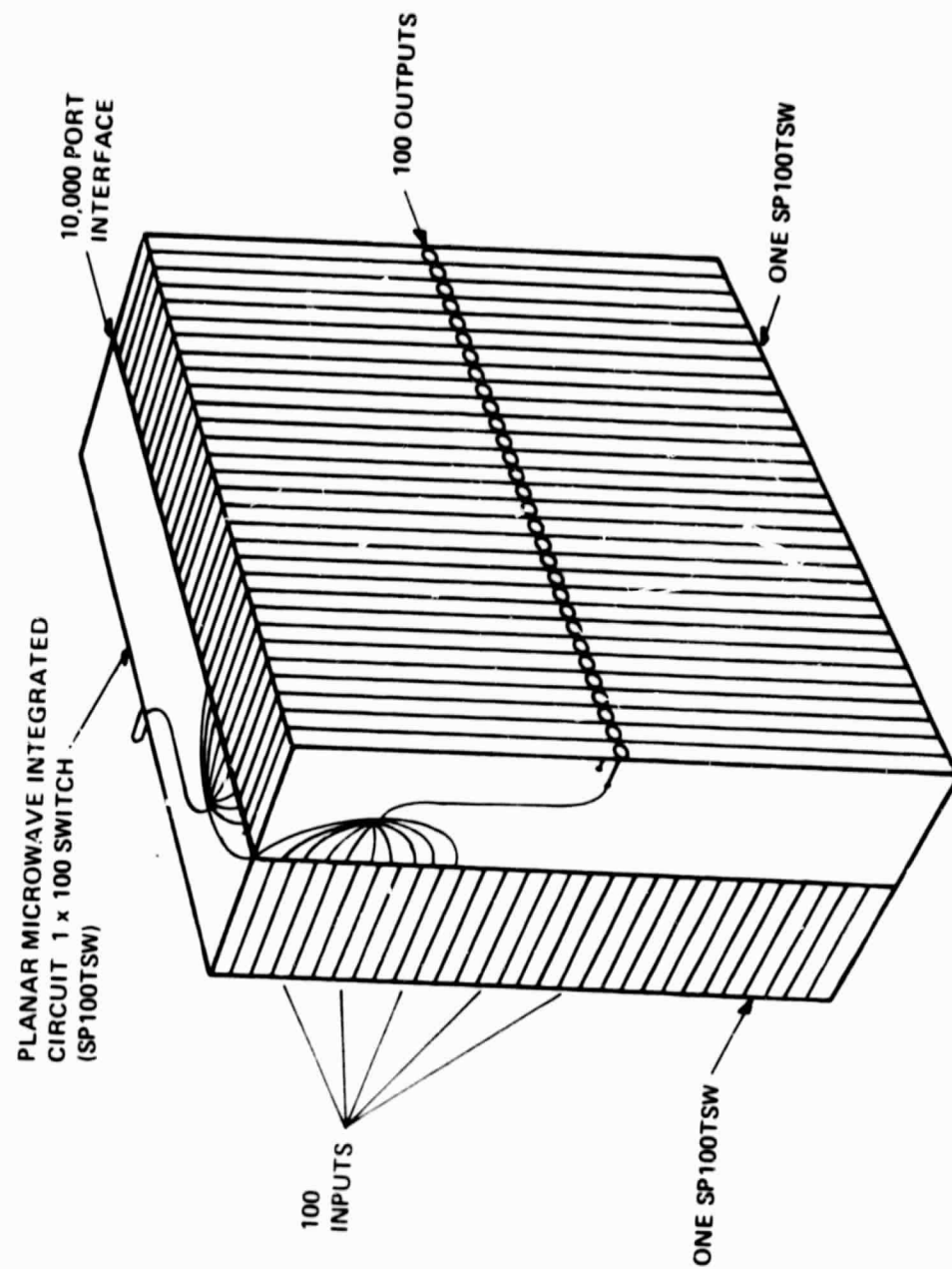


Figure 5-9. Assemblage of SP100T SWs to Form a 100 x 100 Switch

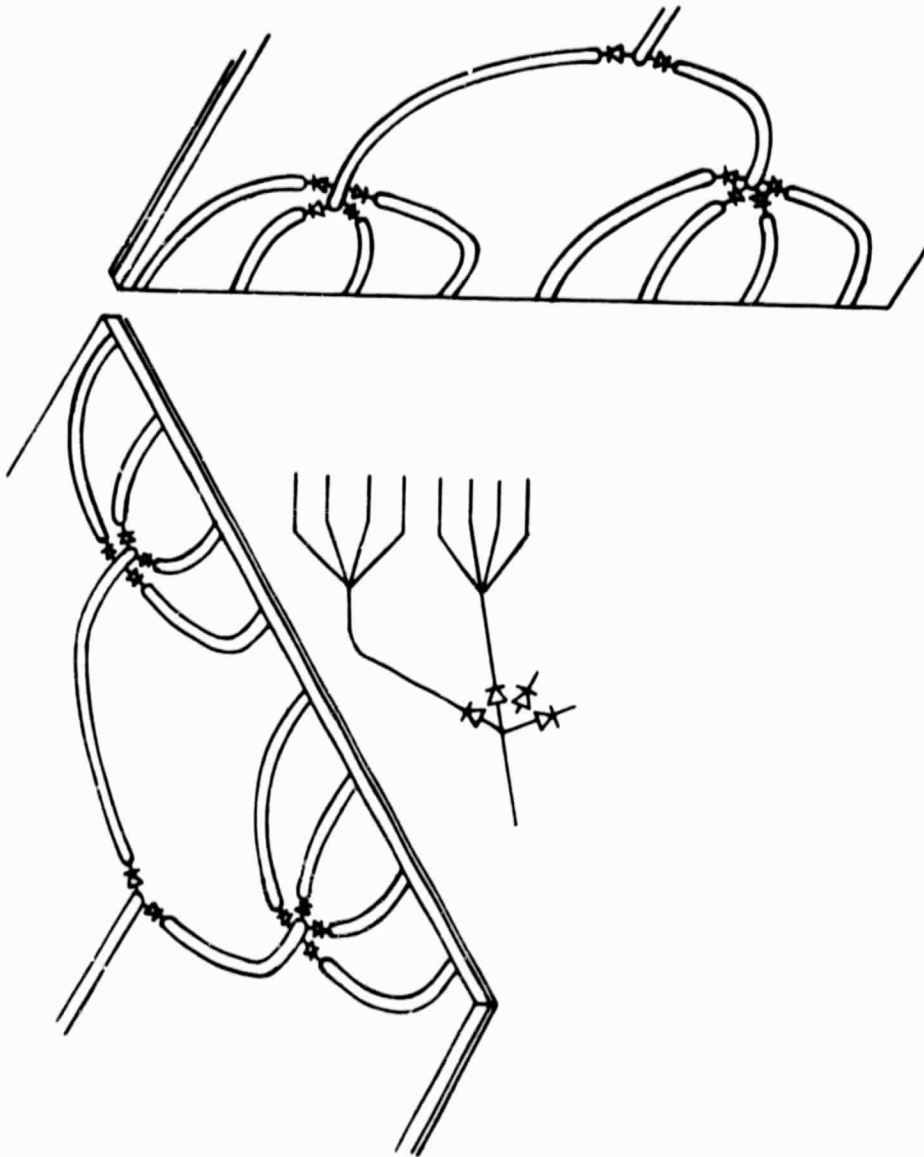


Figure 5-10. Single Pole Multiple Throw Configuration

Each switch has a single input microstrip conductor which ends at a junction. Switching devices, indicated schematically by the symbol for diodes, can direct the signal from the single microstrip to either one of two microstrip lines which leave the junction. The choice depends on how the devices are driven by the logic/driving circuits which are not shown. The signal follows the selected path to a second junction where four devices enable it to leave by any one of four possible output transmission lines. To reach the output on the vertical fan-in board, a similar fan-in procedure is used all the way to the output.

The switch described above has been built by Hughes Aircraft as an 8 x 8 switch using MESFETs. This technique has been suggested by one vendor for constructing a 100 x 100 switch using PIN diodes. Figure 5-11 represents the linear extrapolation of a vendor's quoting rules for small quasi catalog switches to large  $n \times n$  switches and is not binding. This vendor was willing to estimate a figure which is not verifiable as either an upper or a lower bound. The specifications and cost quote are helpful in scoping the cost/complexity of a switch that has never been built.

The figure shows that linearly extrapolating small switch design technology leads to weights, volumes, currents, and power consumption which are too high in the large switch. The conclusion is that a large  $n \times n$  switch should be designed at the outset as a new entity possessed of a unique set of design problems and constraints.

Fan-outs by means of single pole multiple throw switches have certain special performance characteristics. First, there are critical failure points which diminish reliability. Figures 5-7 and 5-8 show that a failure in the input SP10T switch will eliminate at least 10 cross points. If the failure in the input switch were more systematic, all 100 cross points could be eliminated. That is, one entire city would be unable to communicate through the satellite.

The insertion loss through a SPMT fan-out is very favorable. The entire signal, although attenuated, appears at one output. The loss is through  $k$  switches in tandem where  $k$  could be as low as four for a 100 x 100 switch as described. The insertion loss of 10 dB quoted in the hypothetical specification includes both diode and line losses.

If a 1 x 10 module is counted as 10 diode switches with 11 modules to a planar unit, the 100 x 100 switch shown in figure 5-9 has 22,000 switches employing perhaps 44,000 diodes.



# LINEAR EXTRAPOLATION PRESENT SPECIFICATIONS

DIMENSIONS	50" HIGH, 50" WIDE, 12" THICK
WEIGHT	1875 LB
DRIVE CURRENT DRAIN	150 AMP
AVERAGE DRIVE POWER	750 W
AVERAGE DIODE ARM (I)	15 MA
POWER DIV. DRIVE/DIODES	0.8/0.2
SWITCHING SPEED	50 ns
MAX PATH INSERTION LOSS	10 dB
ISOLATION	60 dB
BAND WIDTH	2.5 GHz (NO PROBLEM)
CENTER FREQUENCY	3.0 TO 4.0 GHz PREFERRED
COST @ \$50 PER ARM	\$500,000
ENVIRONMENT	SPACE QUALIFIABLE BUT NOT QUALIFIED
RELIABILITY	METHOD OF CALCULATION NEEDS DEFINITION

Figure 5-11. Performance of a 100 x 100 Cross Bar

### 5.1.3.3 Coupler Fan-Outs

It is not necessary for a fan-out/fan-in structure to be comprised of SPMTs. It can be constructed of hybrid power dividers as suggested in figure 5-12. There still are horizontal and vertical layers as shown in figure 5-9 but a layer contains a different printed circuit. Figure 5-12 shows the makeup of the fan-out circuit utilizing Wilkinson power dividers, which are four-port divider hybrids with the fourth port suppressed at the center of the resistor. The hybrid has the effect of decoupling the two output arms while dividing the input power equally and in phase between them. Each of the output arms then becomes an input arm for the next coupler. The power is successively divided until it appears equally divided at all the output ports. As with the SPMT arrangements, vertical coupler fan-ins are used to complete the topological connections.

The switches are located in the vertical plane between the horizontal fan-outs and vertical fan-ins at the places indicated by the label "diode loc". The switches can be devices other than diodes; a switch is located between every fan-out output and every fan-in input for a total of 10,000 switches.

The power divider architecture has a specific set of advantages and disadvantages. Because the power is divided into  $n$  fractions and appears nominally at  $n$  ports at  $1/n^{\text{th}}$ , the level that it entered the switch, the insertion loss,  $I$ , of an  $n \times n$  crossbar is given by

$$I = 2 \times 10 \log_{10} n'' + S_1 + T_1 \text{ dB}$$

where

$$S_1 = \text{switch loss dB}$$

$$T_1 = \text{transmission line loss dB}$$

$$n'' = 2^k: n \leq 2^k, k \text{ min.}$$

For example, a  $100 \times 100$  switch would have an attenuation of at least 45 dB from input to output port. Placement of the microwave switch at the output of a receiver chain, where both the signal level and the signal-to-noise ratio (under normal operating conditions) are set at a sufficiently high level, overcomes this attenuation.

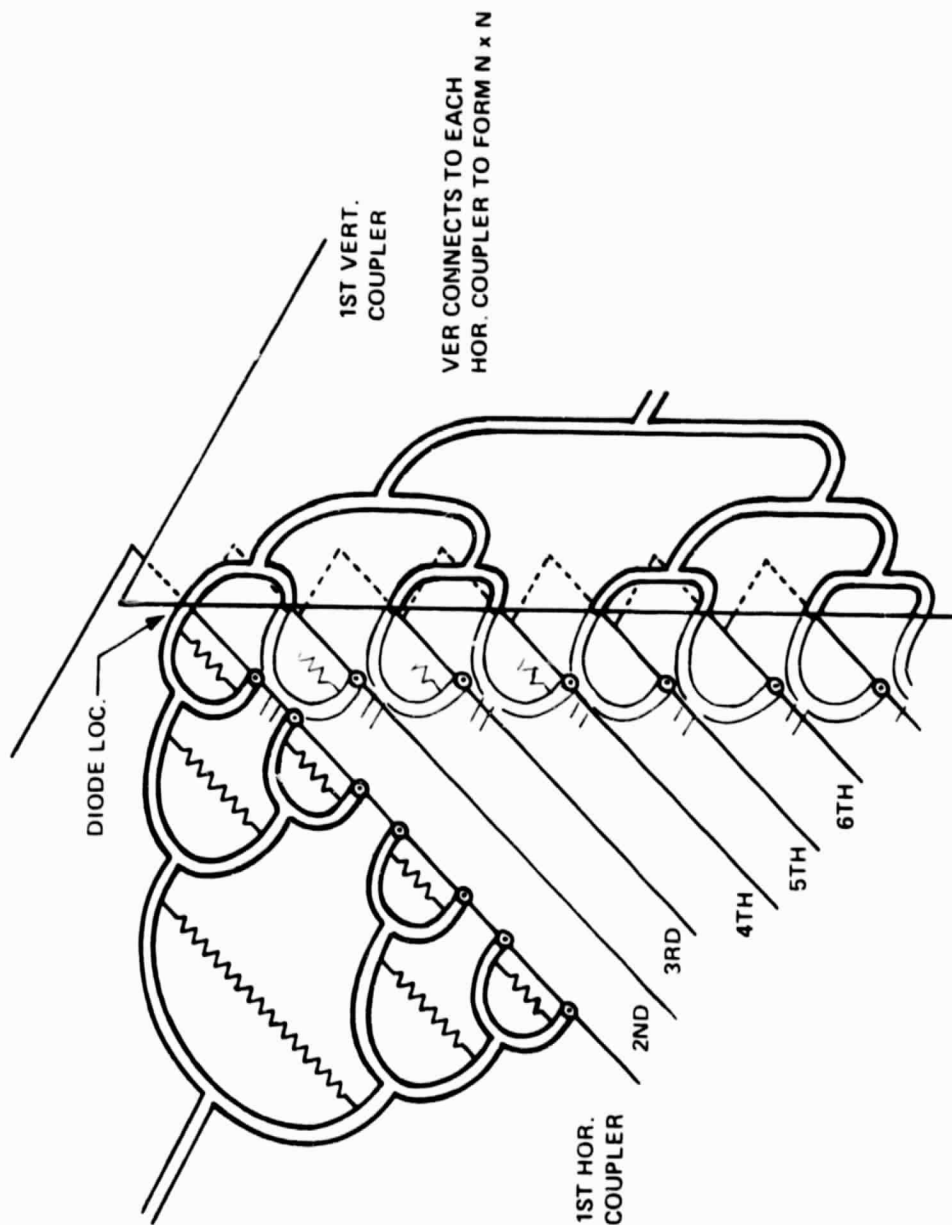


Figure 5-12. Power Divider Configuration

The power divider structures forming the horizontal layers and vertical layers of the fan-outs and fan-ins weigh as much as the corresponding units in the preceding SPMT arrangement. However, there is an additional central section not shown in figure 5-12, which houses the switches. The external arrangement is shown in figure 5-13. This center section has two interfaces with the fan-out/fan-in coupler sections and is therefore more difficult to build. It also causes about a 50 percent addition to switch weight and volume. Offsetting these two disadvantages is improved reliability. A single switch failure loses only the specific cross point involved. Implementation of redundancy to overcome the loss of a switch point must be external to the switch structure and is possibly awkward. Another advantage of the power divider fanned structure is its inherent capability to provide a one to many mode of operation since any input is simultaneously available to each output fan-in.

The Wilkinson power divider form of switch has been built under a study for Intelsat by Thompson-CSF described in *Microwaves* magazine [Microwaves, 1977], shown in figure 5-14. The data supplied in the news release are incomplete because input standing wave ratio (SWR) is not covered. There is potentially a high input SWR problem because all the switches but one in a fan-out are open circuited. Microwave power reflects back and is not absorbed in the loads of the dividers because, for all dividers but one, both arms of the Wilkinson switch have the same SWR and phase. The coupler is balanced and a fan-out then behaves like a fan-in for the reflections which keep adding at the junctions because of the in-phase condition. Almost all of the reflected power can arrive back at the input and cause high SWR. However, the Wilkinson divider is an economical way to realize a 3 dB power division in microstrip. A good compromise to consider is to print all stages of the fan-out but the last using Wilkinson dividers. The last stages should use 90 degree hybrids because these dividers dissipate equal in-phase reflections from their output arms in a load connected to their fourth ports.

#### 5.1.3.4 Rearrangeable Switch Architecture

Figure 5-15 shows an example of a rearrangeable switching network. The elements of this network consist of transfer switches or beta elements developed and employed by Y. Ito, M. Kyogoku, and M. Yamaguchi of KDD in Japan [Ito, 1975; Yamaguchi, 1975] to produce an 8 x 8 switch. A beta element (switch) is one with two inputs A, B, and two outputs, A', B', and two possible states as shown:

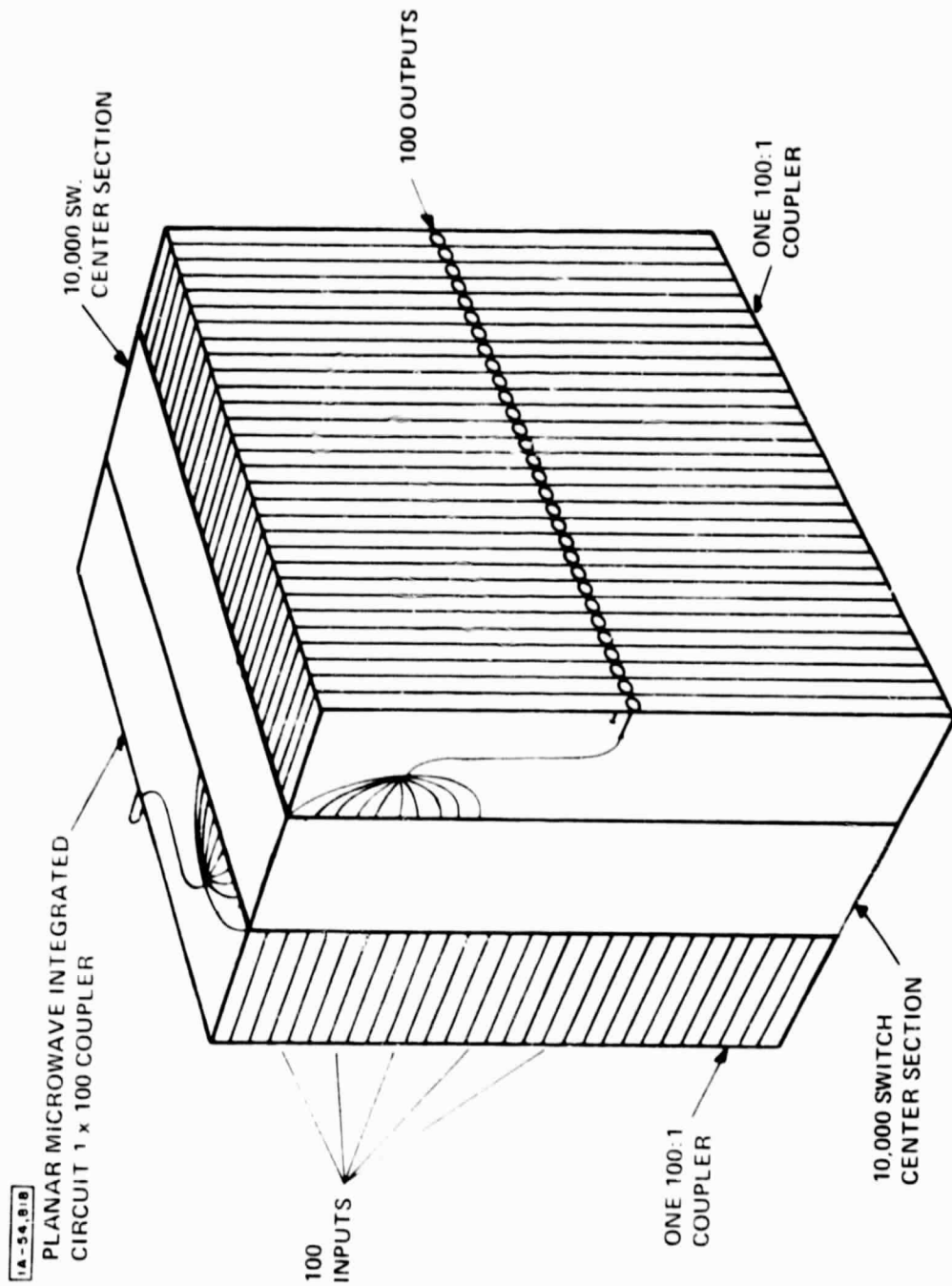


Figure 5-13. Coupler Matrix External View

**New microwave ideas  
revealed in Paris**

Engineers and scientists from around the world converged on Paris, between September 4-8, for the 8th European Microwave Conference. **Papers were presented on blanketing antenna theory and design, advances in semiconductor devices, microwave acoustics, traveling-wave tubes, and communication systems.** The summaries of papers that follow are just a sampling of the areas that caught our attention.

- Engineers from Thomson-CSF, Orsay, France, reported a **16 x 16 fast matrix for frequency reuse satellite communication systems**, operating in the 4-to-6, and 11-to-14 GHz bands. The 256 switches, between the rows and columns of the matrix, are composed of thin-film PIN diodes. In a 3.7-to-4.2 GHz experimental model, a switching speed of less than 50 ns was achieved. Theoretical insertion losses due to dividing and recombining by 16, the switch, and circuit losses, were 24 dB, 0.8 dB, and 2.5 dB, respectively. **Isolation between ways was reported greater than 50 dB and the third-order intermodulation, less than 45 dB.** With an output power of -15 dBm, the total phase shift was less than three degrees. The matrix bias consumed about 8.5 W.

Figure 5-14. Microwaves News Release

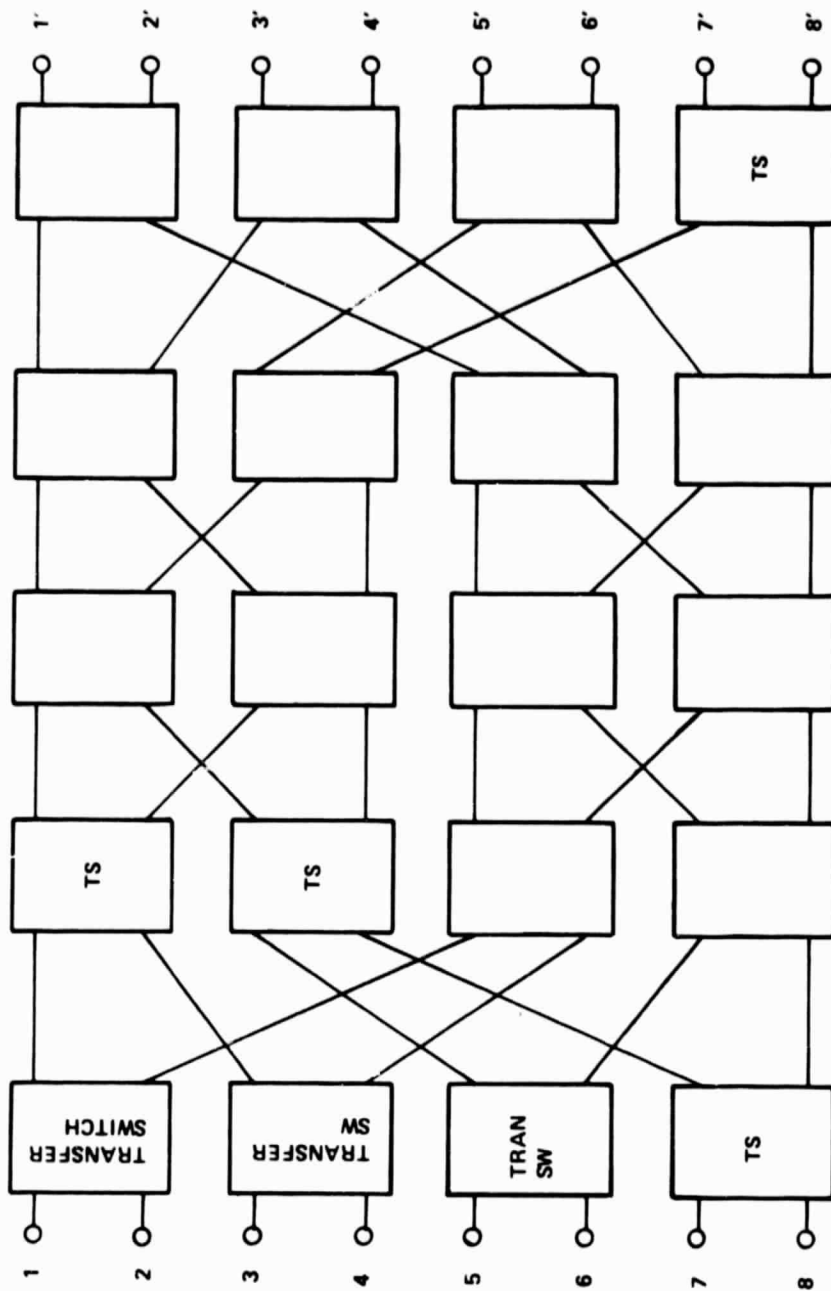


Figure 5-15. Rearrangeable Switch Configuration

### State 1

A connects to A'

B connects to B'

### State 2

A connects to B'

B connects to A'

A network using beta elements, connected as in figure 5-15, can make a connection between any input port on the left and output port on the right under certain conditions and has fewer switches than a switch constructed as a single cross bar (see figure 5-5).

The interconnection of two metropolitan areas during a slot interval is a "call" in the terminology of V. E. Beneš [Beneš, 1965]. To paraphrase Beneš, the calls in progress in a connecting network do not usually arise in a predetermined time sequence. Requests for connection (new calls) and terminations of connection (hangups) occur more or less at random. The performance of a network is measured by the fraction of requested connections that cannot be completed. This is called the probability of blocking. If the structure of a network has too few switches and is thus too simple, it will be inexpensive to build but may have a high probability of blocking. If a network is extensive, it may have low blocking probability but be costly to build and difficult to control.

C. Clos has studied a whole class of networks that are non-blocking because it is always possible to connect an idle pair of terminals without disturbing calls already in progress [Clos, 1953]. Such a network is called non blocking in the strict sense because it has no blocking states whatever. A network may have potential blocking (if paths for new calls are wrongly chosen) which can be avoided by carefully following a set of route assignment rules for each new call. Such switches are classified as non blocking in the wide sense. A network is called rearrangeable if for any set of calls in progress, and any pairs of idle terminals, the existing calls can be reassigned new routes (if necessary) so as to make it possible to connect the idle pair. This is the category of the network in figure 5-15.

The beta element can be built using switching devices; TRW has taken advantage of the lower parts count to produce a 4 x 4 matrix for TDRSS/WESTAR [Salcedo, 1977]. The reconfigurable switch is suitable for small numbers of ports and situations in which all slots have the same duration. However, even for a small switch such as a 4 x 4, the complexity of the algorithm for reassigning routes



as port interconnection demand changes is very evident. For a 100 x 100 switch, the computer load for slot to slot changes would be prohibitively high in computation and memory.

Overriding the computer cost to support a rearrangeable switch is the fact that slots are required to be unequal and connection changes must occur at different times for efficient system operation. Cities bursting information during a long slot would be interrupted while alternate routes through a rearrangeable switch are set up for them so new calls can be accommodated. The random interruptions are not acceptable. They represent a slowdown of information transfer and a rise in overhead costs. Interruption is a step back in the direction of making all slots have the same duration which reduces system revenue because of dead time in the slots.

One more disadvantage of the rearrangeable switch is poor reliability score because of the critical failure point existing at the input. If the devices in the input beta element fail, a city is completely disconnected from the communications network. If beta elements interior to the switch fail, then rerouting algorithms needed to accommodate new calls fail. Therefore, interior failures are also critical.

For the reasons given above, rearrangeable architecture is not the most desirable method of serving a TDMA satellite communications network.

#### 5.1.3.5 New Architectural Concepts

Cross bar architecture is less dependent on layering than fanned architecture. As seen from figure 5-5, cross bar architecture can be accomplished in two layers if good feed throughs can be built to interconnect them. Two types of new switch architecture are described which use the feed through concept. Although development work needs to be undertaken in this area to validate these techniques, specifically broadbanding, they present approaches for overcoming some of the problems posed by implementing large switch matrices using previously discussed architectures.

Feed Throughs. Feed throughs have been awkward and difficult to implement at microwave frequencies in printed circuit media. This is because the feed through is an infinitesimal antenna capable of launching a wave in the printed circuit board dielectric. The fields of this antenna are approximations for the radiation of a feed through in microstrip. They are given in MKS units as

$$H_{\phi} = j k I_m L \exp(-j k r) \sin \theta / 4\pi r$$

$$E_{\theta} = j k v \mu I_m L \exp(-j k r) \sin \theta / 4\pi r$$

where  $\theta$  = the angle made with feed through conductor

$\phi$  = azimuth angle

$$k = 2\pi/\lambda$$

$v$  = velocity of light in medium

$\mu$  = electromagnetic permeability

$I_m$  = maximum current in feed through

$L$  = length of feed through in meters

$r$  = radial distance to field point from feed through

$\lambda$  = wavelength of the carrier frequency

$f$  = carrier frequency

$$\omega = 2\pi f$$

$H$  = magnetic field

$E$  = electric field

$$Z_0 = \omega\mu/k$$

The rate of flow of radiated energy is given by the average Poynting vector as

$$P_{av} = 1/2 (E \times H^*)$$

$$P_{av} = Z_0 k^2 I_m^2 \sin^2 \theta / 32\pi^2 r^2$$

When  $P_{av}$  is integrated over a sphere, the total radiated power,  $W_t$  is found to be

$$W_t = 40\pi^2 I_m^2 L^2 / \lambda^2$$

Suppose a 1 mW signal is transferred via a 0.010 inch long feed through which is matched to its output line impedance of say, 50 ohms. Then if a standard level signal of 1 mW is used, the peak current in the feed through will be 0.00632 A. Using the formula just derived for total radiated power, we find

$$W_t = 1.0 \times 10^{-6} \text{ W}.$$

This is a level 30 dB below the specified input signal level of 1 mW so that little power is lost, but it represents an isolation level which may not be high enough to meet the criterion of S/C + n in dB since the switching devices themselves will couple unwanted signals in the outgoing transmission lines. It will therefore be necessary to use shielded feed throughs. These are shown in figure 5-16, which is a first attempt to design a physical cross bar switch. Two one ounce copper microstrip center conductors are connected via a metal pin in the center of a plated through hole. The added shielding needed is provided by the metal collars shown.

Tandem Switch Cross Bar. Two cross points which indicate the construction of the cross bar switch are shown in figure 5-16. A horizontal signal input line is shown entering from the left in the upper layer (drawn with solid lines). The output transmission lines, shown dotted, conduct switched signals downward in the figure. The entering signal approaches a single pole double throw junction at the cross point. If the signal is not to be connected to the output line, the signal is connected to the upper microstrip line in the junction by forward biasing the diode shown. The signal then continues to the next cross point.

If the signal were routed downward, the upper diode would be back biased and the lower diode at the junction forward biased. This would direct the signal to the shorter alternative line in the upper layer which is connected to the first feed through. The signal would proceed to the transmission line in the lower layer which is shown dotted. Another switch would connect the lower layer line to the main part of the output line. A third switch would disconnect the preceding part of the output line in favor of the line from the feed through.

The particular cross bar shown in figure 5-16 has several disadvantages. The first is that four switches must operate in cooperation at every cross point. This requires 40,000 switches for a 100 x 100 cross bar. The second disadvantage is that of reliability. Since a large number of switches must operate in tandem, the failure of any one of them destroys the path. A city would be totally disconnected from the rest of the cities on that path. In addition, the insertion losses can be prohibitively high



for paths where more than 50 cross points must be traversed. Thus, the tandem switch cross bar of figure 5-16 appears unsuitable.

Coupler Cross Bar. The principles of this form of cross bar are illustrated in figure 5-17 which uses an asymmetric strip line configuration. A similar realization exists for microstrip which is superior because of greater decoupling between input and output lines and more convenience in mounting active elements.

Figure 5-17 shows a cross point where an input signal is conducted by a horizontal transmission line from left to right. A portion of the signal either will or will not emerge downward on the vertical output transmission line. In either case, it will be available at the next cross point to the right (not shown) because directional couplers are used. As the signal traverses the cross point region in the figure, it encounters a hybrid directional coupler which couples a certain fraction,  $C$ , of the power onto a coupler transmission line.  $C$  is a decimal less than 1 so that if the power on the input line is 1 mW, the power in the coupler transmission line is

$$1 \times C = C \text{ mW}$$

$$C < 1.$$

The  $C$  milliwatts of coupled power then propagate to the left on the coupler transmission line because this transverse electromagnetic (TEM) coupler is contradirectional. Next, the coupler transmission line bends away from the input line and is thereby decoupled from it.

There is a switch in the decoupled region of the coupler transmission line. If this switch is closed, the coupled signal continues around the bend and then enters a second directional coupler region. The second directional coupler couples the signal power to the output line which goes to the destination switch port. If the switch in the coupler transmission line is open, the microwave signal is reflected into the load at the other end of the first coupler, and the signal does not couple to the output line. The ratio of the output signal with coupler switch closed to output signal with coupler switch open is the usual ratio of switch performance  $A = I_C/I_O$ , since the action of the second coupling coefficient is the same in both cases.

There are really two coupler coefficients,  $C$  for coupling to the input line, and  $D$  for coupling to the output line. These values are to be selected to operate the switch efficiently. The reason for using couplers to implement a cross bar is to reduce the number

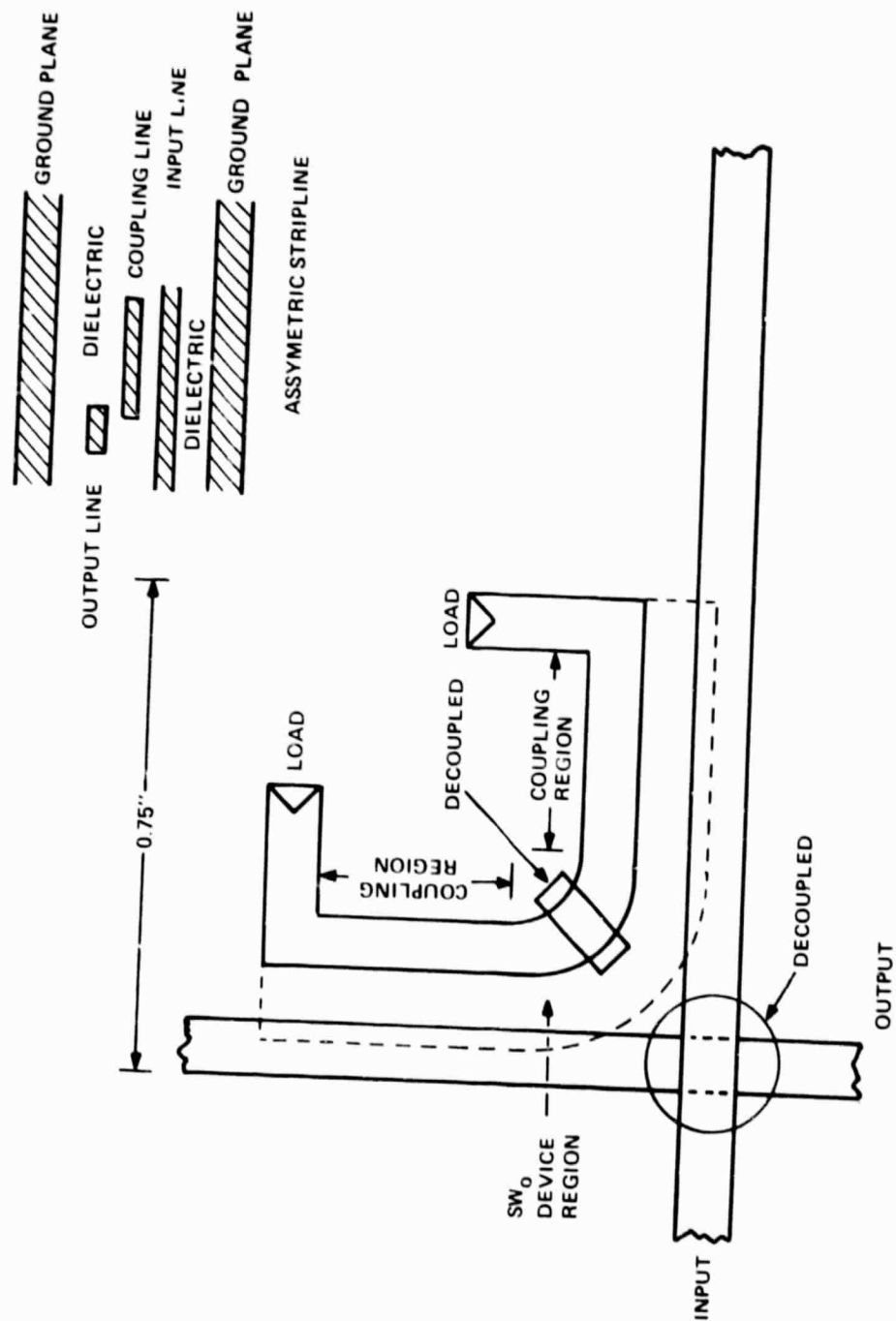


Figure 5-17. Coupler Cross Bar Realization

of switches per cross point, to make the path insertion loss relatively insensitive to the number of cross points traversed, and to greatly increase reliability. This configuration also provides a one to many broadcast capability without special implementation because the signal on the input line can be coupled to any number of output lines simultaneously. It is quite clear that only one switch is used per cross point. This switch may employ amplification, if that is necessary, by providing a switch in series with an amplifier, or by using an amplifier as a switch.

The signal encounters its major attenuation in a coupler cross bar only once, when traversing two couplers and a switch in transferring from an input line to an output line. The losses in passing other couplers are held to minimum levels. Figure 5-18 shows a coupler cross bar with a multiplicity of cross points. The couplers are arranged to couple signals entering from the left and traveling horizontally to vertical transmission lines on which they would proceed upward towards output ports. Input and output lines are terminated to provide a matched system. The switches/amplifiers are not shown. The following calculations indicate the performance of a 100 x 100 switch.

The question to be answered is, "What constant coupling value of the first coupler will provide the largest extracted signal for the 100th output line?" Call the coupling coefficient of the first coupler of each cross point  $C$ . Then the power coupled off at the first cross point,  $P_1$ , is given by

$$P_1 = 1 \times C \text{ mW}$$

for a 1 mW input level. The power remaining after the signal passes the first coupler is the figure needed. It is

$$P_{r1} = (1 - C)$$

Before arriving at the 100th coupler, the signal will have passed 99 couplers and will have a remaining power of  $P_{r99}$  where

$$P_{r99} = (1 - C)^{99}$$

The last coupler will couple a fraction of power  $C$  from the power reaching it. The power extracted for the 100th line,  $P_{100}$ , will be

$$P_{100} = C \times (1 - C)^{99}$$

Let  $N = n - 1$ . The general form of the equation for an  $n \times n$  switch is  $P_n = C \times (1 - C)^N$ . This is differentiated as a function of  $C$  to maximize the power  $P_n$ .

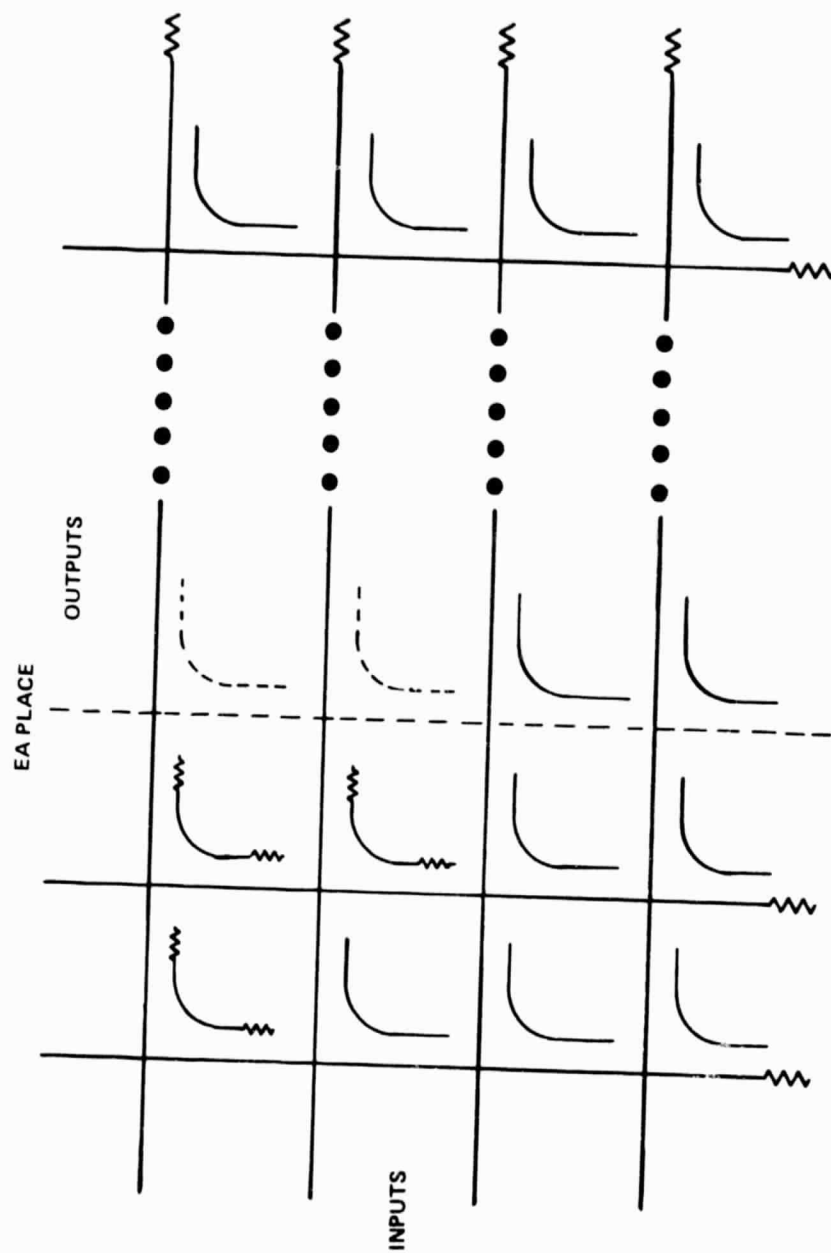


Figure 5-18. Cross Bar with Multiple Cross Points



$$\frac{d}{dc} (P_n) = (1 - c)^N - Nc(1 - c)^{N-1}$$

or

$$1 - c - Nc = 0$$

$$1 - (N + 1) c = 0$$

$$c = 1/(N + 1)$$

$$c = 1/n$$

Further differentiation shows that this value is a maximum. To maximize the signal extracted for the nth line, the optimum coupler value is given by

$$C_{opt} = 10 \log_{10} n$$

A number of values of C were used to calculate the level of  $P_{100}$ . The results are shown in table 5-2.

Table 5-2

Power Extracted for 100th Line

<u>Coupling Level, C. (dB)</u>	<u>Power Extracted (dB)</u>
-10	-55.3
-15	-29.0
-20	-24.3
-30	-30.4

It should be noted that the most power is extracted for a coupler value of 20 dB, equal to  $10 \log_{10}$  of the number of ports.

In table 5-2 the power extracted for the 100th line for C, -20 dB is at the level of -24.3 dB. The power extracted for the first line is -20.0 dB. This means that the variation in power extracted for 100 output lines is only 4.3 dB, a very reasonable figure.

Every cross point has a first coupler for extraction of coupling value C, a switch/amplifier, and a second coupler for inserting the signal on the output line of coupling value D. Consequently, the signal inserted on the output line has the level,  $S(d)$ , given by

$$S = D \times C \times (1 - C)^n$$

The signal then passes a number of such second couplers on its way to the output, and each one takes a small amount of power  $(1 - D)$  from it. If the signal were to pass  $k$  of the second couplers before reaching the output, then the level of the output signal  $S_0$  would be

$$S_0 = D \times (1 - D)^k \times C \times (1 - C)^n$$

The values of  $C$  and  $n$  are now fixed and it is the value of  $D$  that needs to be optimized. The same process of differentiation used before is applied to  $D$ . We note that  $S_0$  is the same function of  $D$  that  $P_n$  was of  $C$  except for a constant  $(C \times (1 - C)^n)$ . The answer is

$$D = 1/k$$

It seems reasonable to optimize again for the lowest input signal row,  $n$  because there is no way to optimize the couplers above a given row,  $r$ , without having a non optimum arrangement for the rows below it. Therefore, we optimize for the lowest row and make  $D = 1/n$ .

With the rules just developed, the 100 x 100 cross bar switch will have a maximum insertion loss,  $L$ , of

$$L = 10 \log_{10} (C \times D) + 2 \times 4.3 \text{ dB}$$

plus line losses. If  $C = D = 0.01$ , as is likely, the insertion loss will probably be 50 dB, an acceptable value, if amplification is not used at the switching points. It is possible that amplification could be used with one form of switching device.

It is thus demonstrated that it is possible to physically construct the switch as a cross bar with acceptable insertion loss and high reliability. The failure of a passive switch cross point loses only that cross point. Neither a row nor a column is eliminated, nor are other cross points affected if the switch fails open. If it fails closed, then the signal to cross talk on one output line degrades to unity if a switch is passive. If amplification is used at a cross point for switching, then no degradation will occur on any failure other than the loss of a cross point.

Figure 5-17 shows a cross point which occupies a square 0.75 inch on a side. This is considered to be a conservative estimate. The number is more likely to be 0.50 inch on a side. However, the 0.75 inch dimension is retained and a size is estimated for the switch as a whole. The calculation is as follows:

1. Area for a cross point  $0.5625 \text{ in}^2$  (based on a dielectric constant of 10)
2. 10,000 cross points require  $5625 \text{ in}^2$
3.  $5625 \text{ in}^2 =$  a square 6.25 ft on a side
4. The square would be no more than 0.5 in. thick
5. Total volume =  $2812.5 \text{ in}^3 = 1.628 \text{ ft}^3$
6. This volume is 9.4 percent of the size of the vendor switch discussed earlier, a reduction by a factor of 10.

A 6.25 foot square is too unwieldy to mount in a satellite, instead it will be cut into rectangles each of which is a smaller cross bar switch. The switches could then be connected by RF coaxial assemblies. Suppose, for example, that the switch was cut into cards 7.5 inches wide by 15.0 inches long. Such cards would each hold 200 switches. If a rack drawer were able to hold 20 cards, then three rack drawers 10 inches high, 18 inches wide, and 15 inches deep would be able to house the entire switch with room to spare for 10 extra cards.

The dimensions chosen are for illustrative purposes only. A tradeoff is needed as part of the design effort to find the best partitioning of the physical cross bar. The matter has more important aspects than partitioning of the larger matrix merely for the purpose of packaging. Clos has shown how, while keeping the strictly non blocking character of the switch, to reduce the number of cross points required by partitioning a cross bar matrix into cross bar submatrices [Clos, 1953]. Clos points out that it is possible to reduce a  $100 \times 100$  switch normally requiring 10,000 cross points to one that only requires 5700. Full advantage should be taken of Clos's partitioning and interconnecting technique. For a 3-stage network, as illustrated in figure 5-19, this requires a number of cross points  $C$  given by  $C = 6 N^{3/2} - 3 N$  for an  $n \times n$  switch. This number is less than  $N^2$  as shown for  $N \geq 36$  in the following illustration.

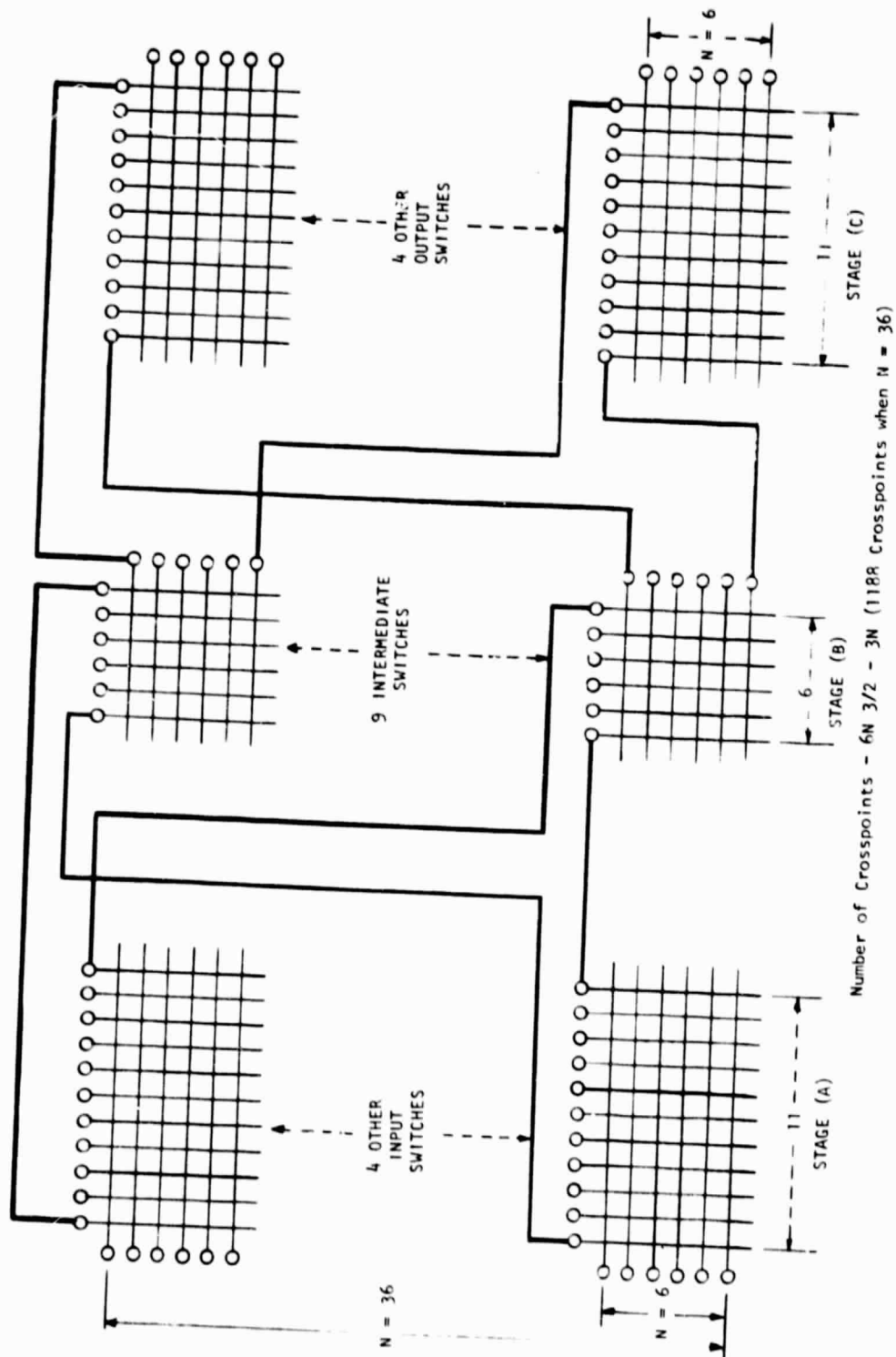


Figure 5-19. Three Stage Connecting Network

Cross Points for an N x N Switch

N	$N^2$	$6N^{3/2} - 3N$
4	16	36
25	625	675
36	1296	1188
64	4096	2880
100	10000	5700

The use of directional couplers makes it possible to build the microwave switch as a physical cross bar. This form of construction is superior because:

1. The least size, volume, and weight are required.
2. Partitioning into smaller cards makes their manufacture simpler and less costly.
3. Partitioning makes it possible to size the cards (switch submatrices) and to connect them as Clos networks and so reduce the number of cross points appreciably.
4. Reliability is enhanced because the switch can be designed so that failure of a device means at most the loss of one cross point.
5. The largest part of the insertion loss is taken once on passing through two couplers. The rest of the insertion loss is small and total insertion loss is acceptable for a modest input power per line of 1 mW.
6. A one to many broadcast mode is an inherent capability of the architecture.
7. The system is easy to match, easier than the Wilkinson dividers or the SPMT fan-outs.
8. Redundancy is simpler to implement than with fan-out architecture.

There is a problem in broadbanding the directional couplers over a 2.5 GHz band. Multistage couplers may have to be used, the design of which is given in the literature [Tajima, 1978]. Some effort will be required to maintain the 0.75 inch dimension assigned to a cross point. It is believed that an acceptable solution can be delivered. The many advantages of cross bar architecture based on couplers make it the most promising candidate for realizing a 100 x 100 switch at microwave frequencies.

#### 5.1.3.6 Phased Array Architecture

A phased array might be conceived as a method of providing switching between high traffic areas. In the case considered here, the uplink beam from each city is connected to a downlink phased array from the satellite which scans from city to city and dwells in accordance with the source city frame and slot schedule. All cities transmit simultaneously. For flexibility in designation of source traffic areas and the possibility of switching, phased arrays are also used for the uplinks. The satellite therefore has a large receiving and a large transmitting phased array, each producing many independently scanned beams.

A phased array is an antenna consisting of radiating elements whose beam can be moved even though the aperture is stationary. This is accomplished by using a dedicated phase shifter for each element. The application of a linear progressive phase shift across the aperture tilts the beam in the direction of lagging phase according to well known rules. This obviates the need for moving the aperture mechanically. The antenna beam can move at electronic speeds. Sometimes time delay is needed in addition to phase shift to support bandwidth.

Figure 5-20 is a schematic of a phased array which has been discussed for TDMA transmission from a satellite [Acompa, 1978]. It was used in discussing the relationship of capacity to transponder number and is used here to discuss phased array component requirements.

The method for producing multiple beams with a phased array of  $P$  elements is to split the received (transmitted) signal from each radiator into  $T$  samples and process each sample differently, as suggested in figure 5-20. Suppose all the number one samples from each radiator have dedicated phase shifters programmed to receive signals from a certain direction  $\theta_1$ , and that after being phase shifted these samples are added somehow in a  $P$  to 1 combiner. Then a receiver at the output of that combiner will receive signals with a beam pointing in the direction  $\theta_1$  and having a gain and resolution appropriate to  $P$  elements. There are  $T$  combiners and

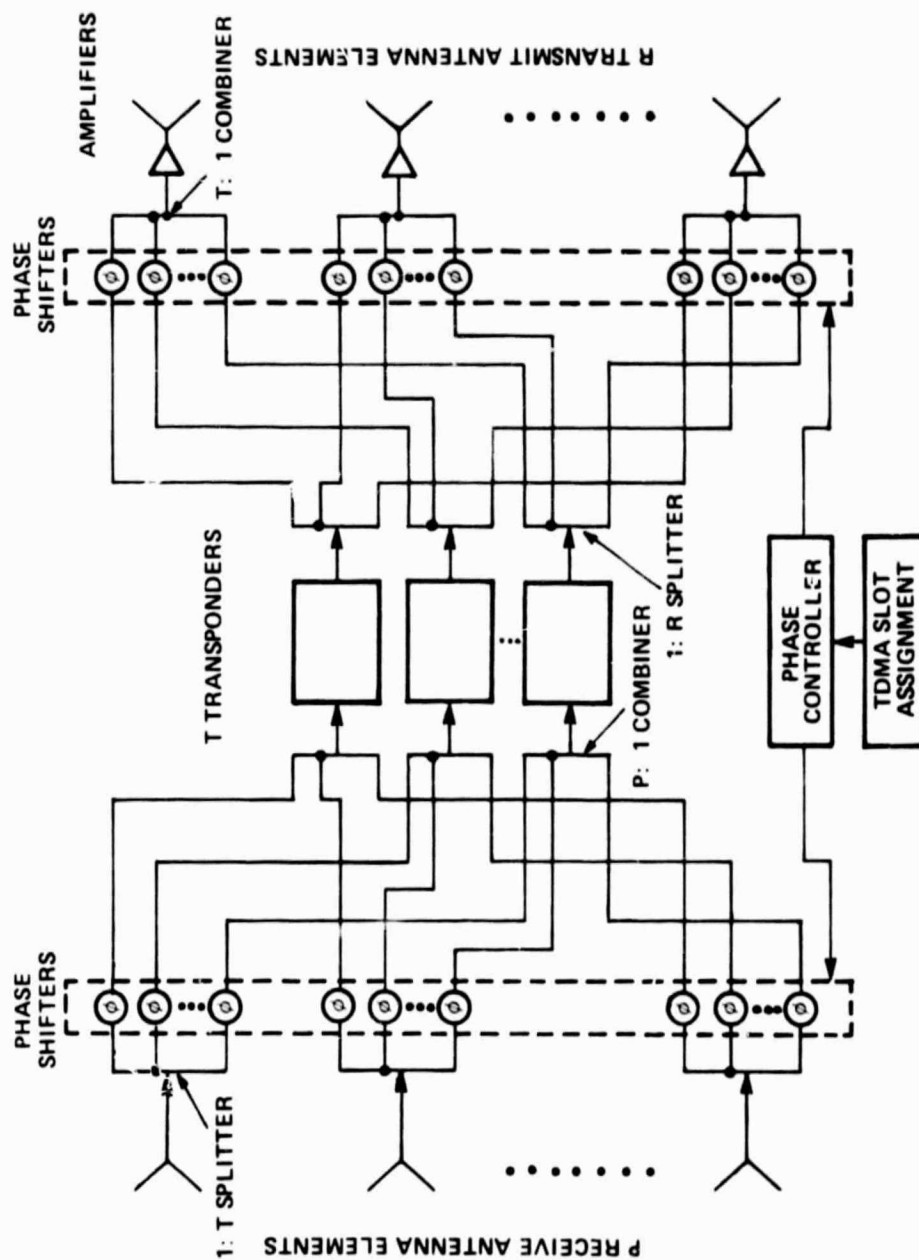


Figure 5-20. Phased Array for TDMA Scan



T x P phase shifters. (Certain details such as positions of amplifiers are omitted but are not crucial to discussion of the principles.)

There are certain drawbacks to such a scheme for using the phased array as a switch. The first of these is the excessive cost associated with the large number of phase shifters required for the antenna. Even if the antenna is thinned, it can be shown that a 15 foot antenna with 50 beams would require in excess of  $2 \times 10^6$  phase shifters. This is an extraordinary number of phase shifters to buy, mount, power, control, and boost into orbit.

The second problem has to do with microwave circuitry. A number of lossless multiport microwave junctions are needed which add P or T signals to produce a single output. This must be a true vector sum of any set of arbitrary phasor signals. However, microwave junctions do not operate this way. Instead, they make a complex linear transformation on the set of T input signals. The output is not, therefore, the vector sum of the inputs. A true vector sum may require extremely expensive processing, and  $2 \times 10^3$  such processors would be needed.

The third drawback has to do with the output amplifiers shown at the right of figure 5-20. It is desirable to have these amplifiers after the summers so that only one is used per radiator. It is not expected that these amplifiers will have the linearity in the 80's and 90's to add five 500 MHz bands with an acceptably low intermod level. This casts doubt on the ability of the array to perform its communications task.

A phased array has inherent technical difficulties in its summers and amplifiers. It uses an excessively large number of phase shifters and associated control components. These would require unacceptable cost and weight. For these reasons, phased array architecture will require inordinately large investment in research and development to attempt to interconnect a large number of high traffic areas by this technique.

## 5.2 PIN DIODE TECHNOLOGY

Microwave switching by means of diodes is a mature technology which started in November 1955 with the work of [Armistead and Spencer, 1956]. Since then, the switching of microwaves at electronic speeds has supported a small industry whose products include numerous switch configurations, microwave modulators, attenuators, and microwave phase shifters. Two comprehensive books have been written on the subject [Garver, 1976; White, 1977]. This



section covers some diode principles relating to the design of a large microwave switch and assesses the applicability of diode technology for this purpose.

#### 5.2.1 Basic PIN Diode Operation

Figure 5-21 shows a series diode switching circuit. It consists of a diode switch and a biasing circuit. The simplified circuit diagram of an ordinary P - N junction diode is shown in figure 5-22. The junction resistance  $R_j$  depends on the bias supplied to the diode. When the diode is forward biased, high d.c. reduces  $R_j$  to a small value (1 to 5 ohms) and the series-diode switch shown in figure 5-21 is closed. In this state, most of the RF power applied to the input of the switch appears (with small loss) at its output. When the diode is reverse biased,  $R_j$  becomes very high (above 1 kilohm) and the switch opens. In this isolated state only a small portion of the input power appears at the output, mostly via the junction capacitance of the diode, which can be quite small (0.02 to 1 pF).

In ordinary P - N junction diodes, the relation between diode current and junction voltage is non-linear and therefore the junction resistance  $R_j$  defined as  $dv_j/di_j$  is a non-linear function of the forward current.  $C_j$  is also non-linear, depending on the junction voltage. In Schottky-barrier diodes, the non-linear behavior of  $R_j$  is used for mixing; in varactor-diodes the non-linear behaviour of  $C_j$  is used for multiplication. In a switching application non-linear  $R_j$  and  $C_j$  are undesirable because they cause signal distortions.

PIN diodes were developed to overcome those deficiencies in diode switches. A PIN diode consists of three semiconductor layers, a positively doped P layer, a neutral intrinsic I layer, and a negatively doped N layer. The insertion of an undoped I layer decreases the capacitance. This makes the diode a good open circuit when back biased, and reduces undesired non linearities. The construction also provides more area which lowers series resistance and thus loss when the switch is closed.

An important parameter of the PIN diodes is the charge carrier life time  $t_1$  in the intrinsic layer. If a switch is closed and passing current, the charges in the I-layer have time to recombine and the PIN diodes behave as an ordinary P - N junction diode when the RF frequency is low. When the RF frequency is high (above a value  $f_0$ , defined as  $f_0 = 1/2\pi t_1$ ) the PIN diode becomes too sluggish to follow the RF signal and exhibits a linear resistance and capacitance,  $R_1$  and  $C_1$ . A diode switch uses PIN diodes operating

1A-54,803

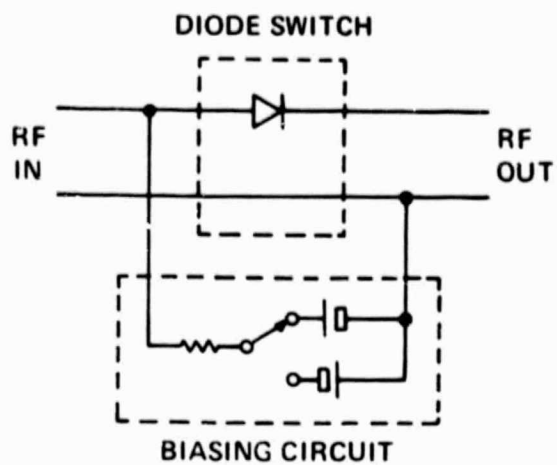


Figure 5-21. Diode Switch

1A-54,804

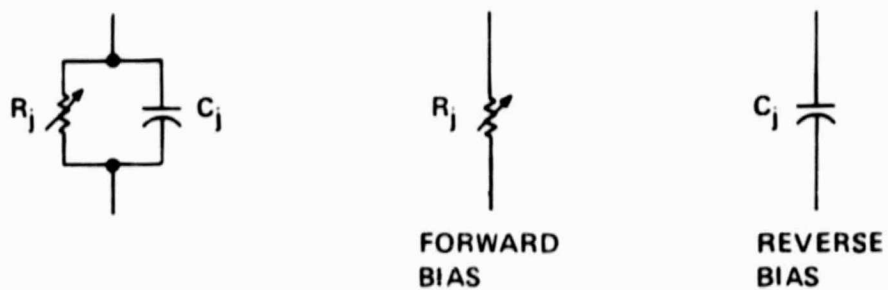


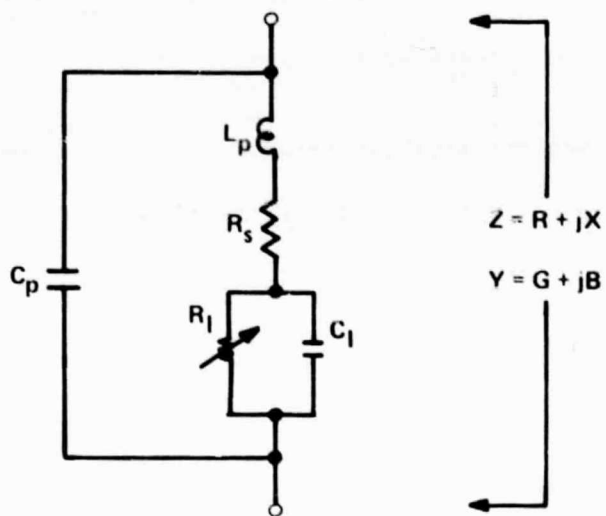
Figure 5-22. Simplified Equivalent Diode Circuit

above  $f_0$ . Depending on the diode,  $t_1$  can vary from 1 to 1,000 nsec. If  $t_1$  is 100 nsec, the RF frequency has to exceed  $f_0 = 1.6$  MHz. In practice, the operating frequency should be much higher than  $f_0$ , about 100 MHz in this example.

The stray reactances associated with the diode packaging affect the high frequency response of a diode switch and should be incorporated in the diode's equivalent circuit diagram, as shown in figure 5-23. Note that  $R_j$  and  $C_j$  are replaced by the linear  $R_I$  and  $C_I$ .  $L_p$  and  $C_p$  are the package lead inductance and capacitance, respectively.  $R_s$  is the resistance to current flow in the diode materials.  $R_I$  is the variable diode resistance. It is inversely proportional to the forward d.c.  $I_{dc}$ , or to a better approximation,  $I_{dc}^x$  where  $x$  is a constant slightly smaller than 1.  $R_I$  is proportional to the thickness and inversely proportional to the area of the intrinsic layer and is inversely proportional to  $t_1$ .  $R_I$  should be minimized because it affects the insertion loss in a series-diode switch and the isolation in a shunt-diode switch.  $R_I$  can vary between 1 and 8 ohms for  $I_{dc} = 10$  mA among different diode types.  $R_I$  can exceed 1 kilohm under reverse bias.

$C_I$  depends on the diode geometry; it is inversely proportional to the thickness and proportional to the area of the intrinsic layer.  $C_I$  should be as small as possible because it affects the isolation in a series-diode switch and the voltage standing wave ratio (VSWR) in a shunt-diode switch.  $C_I$  is between 0.02 and 2 pF.  $R_I$  and  $C_I$  are affected oppositely by the diode I layer geometry, as summarized in table 5-3. If isolation is the prime concern, a series-diode switch should have low capacitance with thick I-layer and small area; a shunt-diode switch, on the other hand, has to exhibit low resistance, which means thin I-layer and large area.

$R_s$  is the semiconductor bulk resistance and unlike  $R_I$  is independent of the forward current.  $R_s$  equals the finite residual diode resistance for high forward bias current.  $R_s$  is between 0.5 and 2 ohms.



$$^a \text{ SERIES} = 10 \log_{10} \left[ \left( \frac{R}{2Z_0} + 1 \right)^2 + \left( \frac{X}{2Z_0} \right)^2 \right]$$

$$^a \text{ SHUNT} = 10 \log_{10} \left[ \left( \frac{G}{2Y_0} + 1 \right)^2 + \left( \frac{B}{2Y_0} \right)^2 \right]$$

Figure 5-23. PIN Diode Equivalent Circuit

Table 5-3

## Diode Parameters and Performance

Diode Parameter	Forward Resistance, $R_f$	Low Capacitance $C_f$	Fast Switching	High Power
I-layer Thickness	thin	thick	thin	thick
I-Layer area	large	small	small	large
Charge Carrier Lifetime	high		low	high

$L_p$  and  $C_p$  are parasitic elements which depend on the diode package, as shown in table 5-4. It is not worth inserting a diode chip with  $C_f = 0.02$  pF in a pill package with  $C_p = 0.2$  pF or a diode chip with  $R_f = 1$  ohm (for a given  $I_{dc}$ ) in a glass package with  $L_p = 2.5$  nH. The influence of the stray  $C_p$  and  $L_p$  on the frequency response of diode switches is too detrimental.

Table 5-4

## Parasitic Elements

Package	$C_p$	$L_p$
Beam lead	very small	very small
Chip with bonding wire/ stripline	0.05 pF	2 x 0.25 nH
Pill	0.2 pF	0.4 nH
Glass	0.1 pF	2.5 nH

An important parameter of the diode switch is switching speed. As shown in table 5-3, higher switching speeds are obtained with thin I-layer and small area diodes. Because a high speed diode has a thin I-layer it also has a small  $R_I$  and a larger  $C_I$ . High speed diodes require small  $t_1$ . A general purpose PIN diode could have  $t_1 = 100$  nsec and a rough switching time of 150 nsec, while a high speed diode has  $t_1 = 15$  nsec and switching time of approximately 10 nsec. Switching speed is inversely proportional to the forward current and proportional to the reverse voltage. It can be reduced to  $t_1/10$  by appropriately shaping the driving waveform, as shown in figure 5-24.

PIN diodes will not be required to handle high power in the satellite TDMA switch. Consequently, more expensive high power diode construction techniques are not required.

#### 5.2.2. Single Pole Single Throw Diode Switches

Series Switch. A diode is inserted in series in a transmission system with characteristic impedance  $Z_0$ , as shown in figure 5-25. The diode impedance is indicated in figure 5-23. It can readily be shown that the attenuation,  $A$ , (defined as the ratio of the power delivered to a matched load without the diode, to the power delivered to the load with the diode in the circuit) is

$$A = \left(1 + \frac{R}{2Z_0}\right)^2 + \left(\frac{X}{2Z_0}\right)^2$$

A forward biased diode should introduce minimum attenuation. The insertion loss is obtained on substituting forward biased diode parameters in the expression. Similarly, the reverse biased diode should introduce maximum attenuation. By substituting the reverse-bias diode impedance in the above expression, the isolation of the switch is obtained.

Figure 5-26 shows the insertion loss and the isolation of a single series-diode switch as a function of frequency, taking into account the diode parasitic reactances.

The insertion loss of a 5 ohms forward-biased diode ( $R_d = R_s + R_I = 5$  ohms) from the above formula is

$$A = 10 \log_{10} \left(1 + \frac{R}{2Z_0}\right)^2 = 0.42 \text{ dB}$$

Note that a small insertion loss requires small  $R$ , which in turn requires a high forward biasing current. However, switching speed is inversely proportional to forward biasing current. A small  $R$

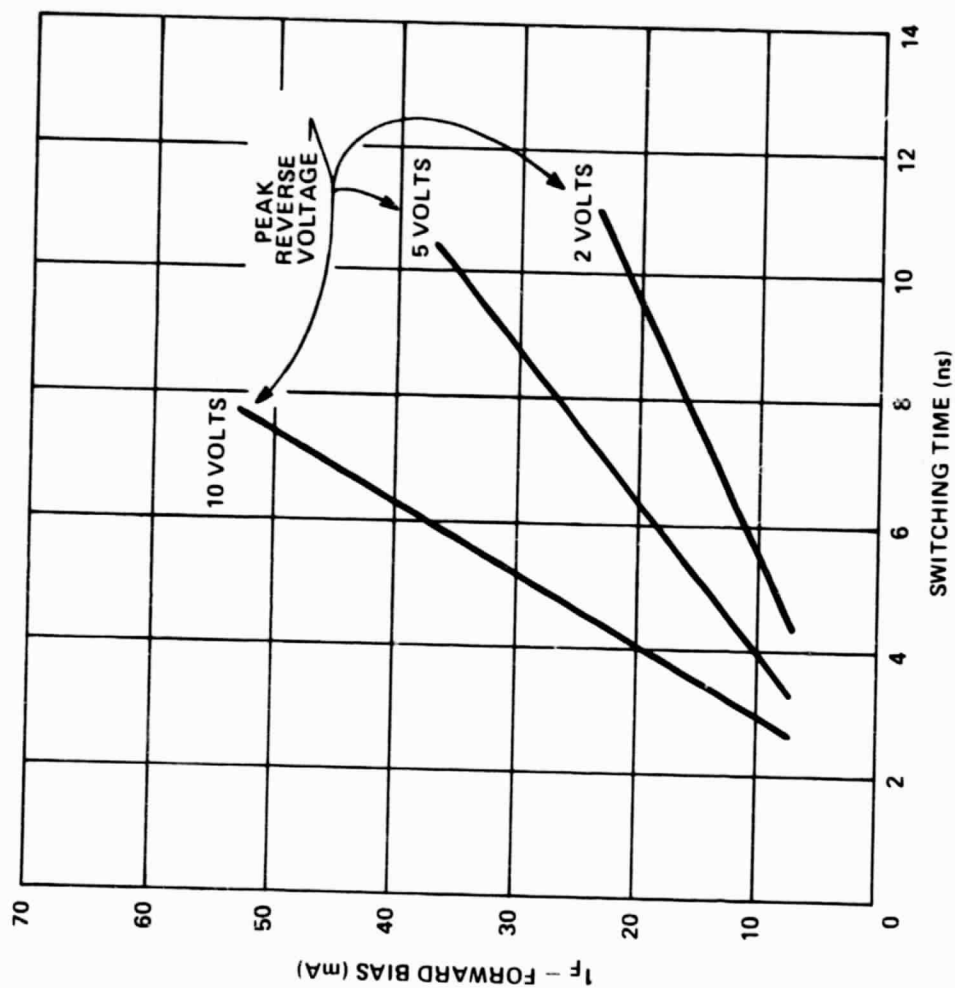
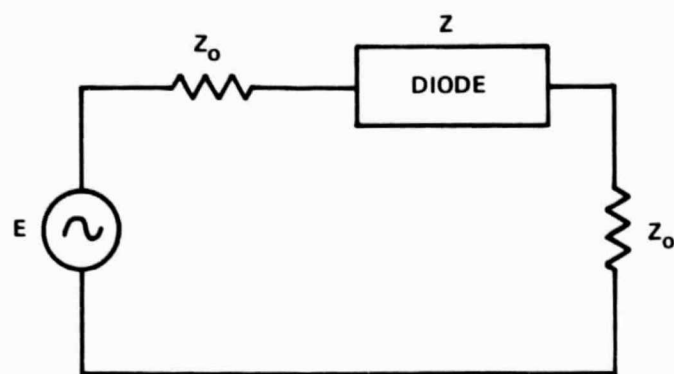
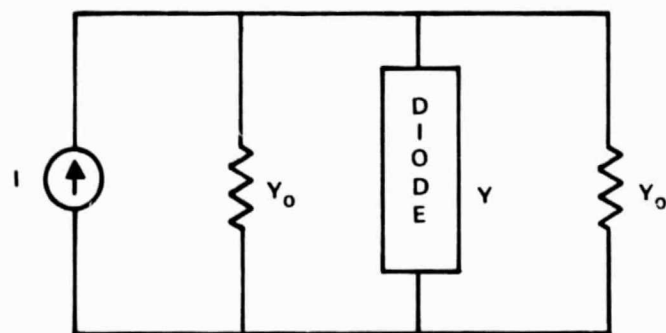


Figure 5-24. Switching Time vs. Forward Bias and Peak Reverse Voltage



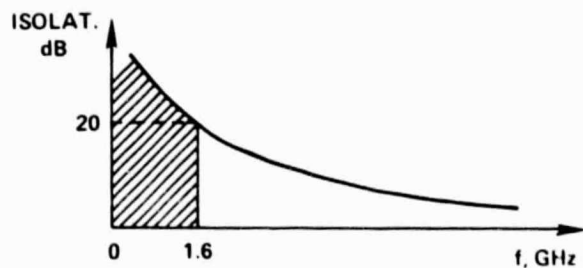
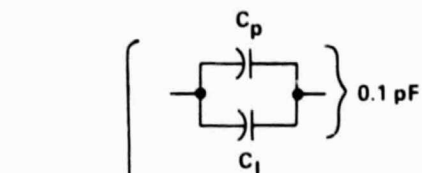
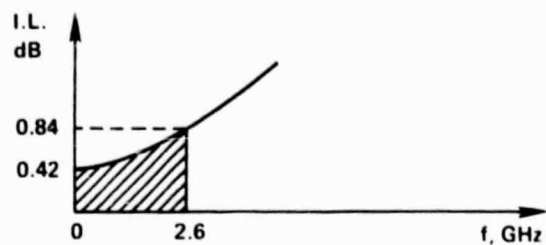
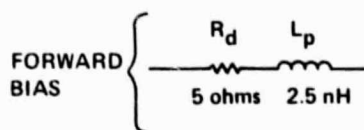
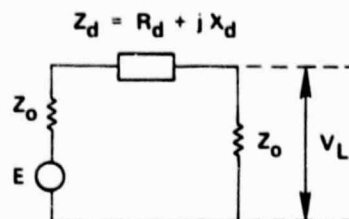
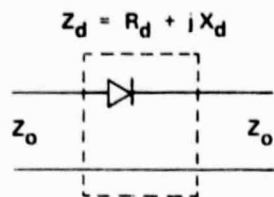
SERIES EQUIVALENT CIRCUIT



SHUNT EQUIVALENT CIRCUIT

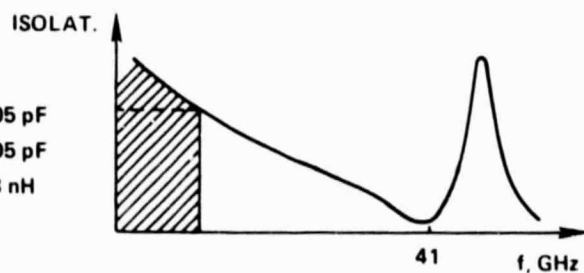
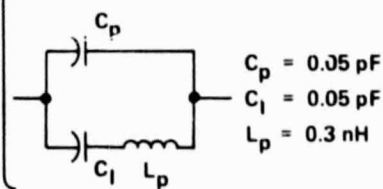
Figure 5-25. PIN Diode Switching Principles





REVERSE BIAS

A MORE ACCURATE REPRESENTATION:



1A-54,801

Figure 5-26. Series Switch Performance

usually is associated with a large  $C_d$  ( $C_d = C_I + C_p$ ), which means a small isolation, as explained below.

The insertion loss depends on  $L_p$  also. Since  $L_p$  of a glass package is rather high, it is not suitable for a switch in the GHz region. Indeed, a 2.5 nH lead inductance will increase the insertion loss to 0.84 dB at 2.6 GHz.

The isolation of the series-diode switch depends on the reverse-bias diode capacity (which is equal to the sum of junction and packaging capacities). A 0.1 pF capacity yields only 20 dB of isolation at 1.6 GHz. A beam-lead diode with reverse capacity of 0.02 pF is the best choice for a high-isolation series-diode switch. The beam-lead diode has a high resistance  $R$ . The total diode resistance ( $R_I + R_S$ ) at 10 mA is 20 ohms, which is an unacceptably high value from an insertion loss point of view. A reasonably low forward resistance can be obtained at higher forward current (6 ohms at 50 mA).

A medium-speed chip offers the best compromise. For instance, Alpha's CSB 7003-01 has a forward resistance of 2.5 ohms at 1 mA, reverse capacitance of over 0.05 pF at -50 volts, and a switching time of about 20 nsec. A reverse capacitance of over 0.05 pF will not provide sufficient isolation in the GHz region and will have to be augmented with additional diodes. Single series diodes alone do not produce sufficient isolation.

Shunt Switch. A diode is inserted in shunt in a transmission system with characteristic admittance  $Y$ , as shown in figure 5-27. It can be shown that the attenuation resulting from the insertion of the diode is:

$$A_s = 10 \log_{10} \left[ \left( \frac{G_d}{2Y_o} + 1 \right)^2 + \left( \frac{B_d}{2Y_o} \right)^2 \right]$$

where

$G_d$  = diode conductance

$B_d$  = diode susceptance

$G_d$  is given in terms of the readily available series diode parameters,  $R_d + j X_d$  as:

$$G_d = \frac{R_d}{(R_d^2 + X_d^2)} \quad B_d = \frac{X_d}{(R_d^2 + X_d^2)}$$

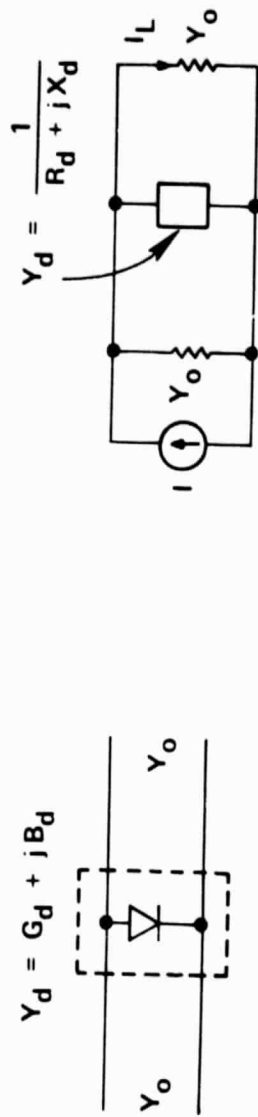


Figure 5-27. Shunt Diode Configuration

The forward biased diode should introduce maximum attenuation. By substituting the forward-bias diode impedance in the above expression, the isolation of the switch can be calculated.

In similar fashion, the reverse biased diode should introduce minimum attenuation. By substituting the reverse-bias diode impedance in the above expression, the insertion loss of the switch is found.

Figure 5-28 shows the insertion loss and the isolation of a shunt-diode switch as a function of frequency, taking into account the diode parasitic reactances.

The isolation of the shunt-diode switch depends on the forward diode resistance. At low frequencies the isolation A, is:

$$A = 20 \log_{10} \left( 1 + \frac{Z_o}{2R_d} \right)$$

For  $Z_o = 50$  and  $R_d = 1$  ohm the isolation is 28 dB, for  $R_d = 5$  ohms it is only 16 dB. The packaging inductance  $L_p$  lowers the isolation for higher frequencies. A diode with  $R_d = 1$  ohm and  $L_p = 0.3$  nH will produce a 20 dB isolation at 1.2 GHz. A value  $R_d = 1$  ohm is an excellent forward resistance, which can be obtained only by driving the diode with a current of over 20 mA. The best candidate for a shunt-diode switch appears to be a diode chip or a stripline/microstrip package. The inductance of the bonding wire can be used as shown in figure 5-29 to form a low-pass filter with the reverse-bias diode capacitance, resulting in a good microwave match and small insertion loss.

### 5.2.3 Multiple Diode Switches

In order to increase isolations to 40 to 50 dB, more diodes are needed. Consider first two shunt diode switches, each diode having a forward-bias resistance of 1 ohm. The isolation per switch is:

$$A = 20 \log_{10} \left( 1 + \frac{Z_o}{2Z_d} \right) = 28 \text{ dB}$$

The isolation of two cascaded switches placed very close to each other is:

$$A = 20 \log_{10} \left( 1 + \frac{Z_o}{2Z_d/2} \right) = 34 \text{ dB}$$

When the distance between the two diodes is  $d = \lambda/4$ , where  $\lambda$  is the wavelength, the second diode short-circuit is transformed into an open circuit at the locations of the first diode by the  $\lambda/4$

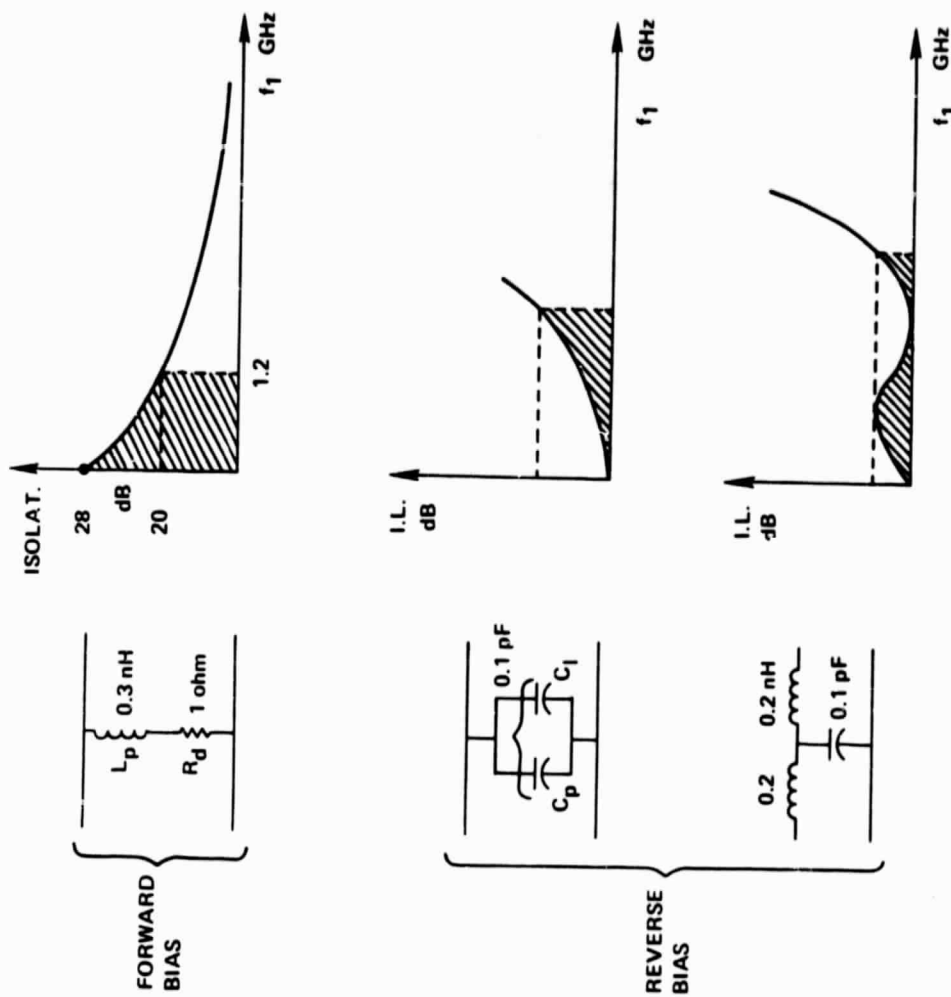


Figure 5-28. Shunt Diode Insertion Loss and Isolation

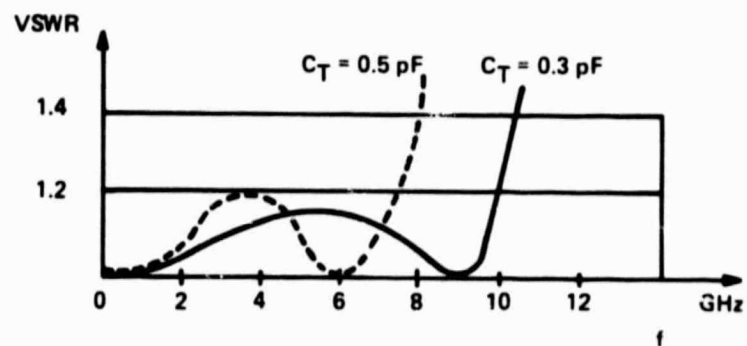
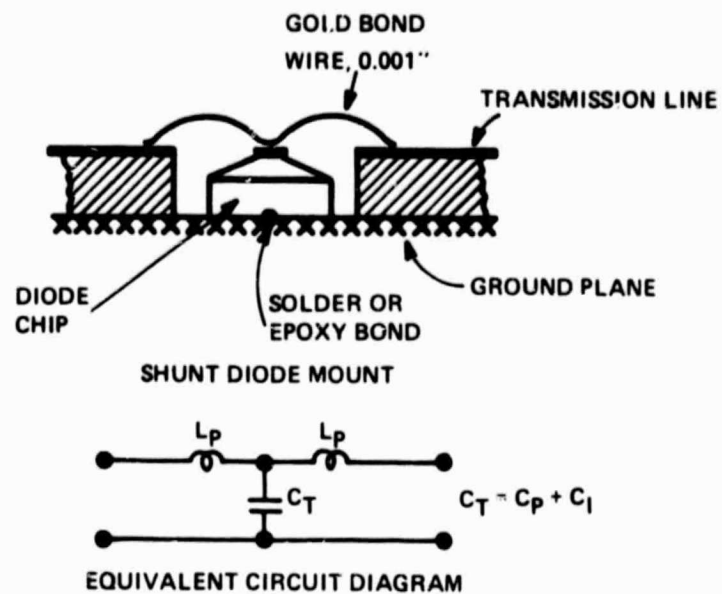


Figure 5-29. Shunt Diode Low Pass Matching

transmission line length, greatly increasing the isolation, as shown in figure 5-30.

This example shows that a rather marginal single-diode switch (28 dB isolation) can be improved greatly by adding a second switch  $\lambda/41$  away, where  $\lambda$  is the wavelength at the center frequency giving an overall isolation higher than 54 dB over a 4:1 frequency band.

Similarly, consider the isolation of two series-diode switches. Each reverse biased diode has a capacity of 0.02 pF and the center frequency is 3.2 GHz. The reactance per diode is  $X_d = 1/\omega C$ .  $X_d = 2500$  ohms, giving an isolation per diode of:

$$10 \log_{10} \left( 1 + \left( \frac{X_d}{Z_0} \right)^2 \right) = 28 \text{ dB}$$

The isolation of the two cascaded switches when placed very close to each other is

$$10 \log_{10} \left( 1 + \left( \frac{2X_d}{Z_0} \right)^2 \right) = 34 \text{ dB}$$

If the distance between the switches is increased to  $1/4$  for the center frequency, the overall isolation increases to 62 dB, as shown in figure 5-31, a situation very analogous to the cascaded shunt-diode switches shown in figure 5-30.

Another commonly used combination is the series-shunt diode switch shown in figure 5-32a. It is a very convenient building block for a single pole N throw switch, also shown in figure 5-32b. The series diodes of each output arm of the switch isolate the shunting diodes from the common input arm.

#### 5.2.4 Switching Speed and Biasing

The various switch configurations have specific biasing problems. A very important characteristic of the microwave switch is its ability to switch rapidly from one state to another. Consider the shunt-diode switch shown in figure 5-33. In the open switch state (isolation), the shunt-diode is forward biased by  $V_f$ . When closed (transmission), the shunt-diode is reversed biased by  $V_r$ . The forward current during the OPEN state is  $I_C$ . To switch the shunt-diode from the CLOSED to OPEN state, charge must be injected into it. To switch from OPEN to CLOSED, this charge has to be removed. The injection and removal of the charge takes time, which is the delay time and switching time shown in figure 5-33. The turn-off time,  $T_{OFF}$ , i.e., the time it takes to switch from transmission to isolation is smaller than the turn-on time,  $T_{ON}$ , which is the time necessary to switch from isolation to

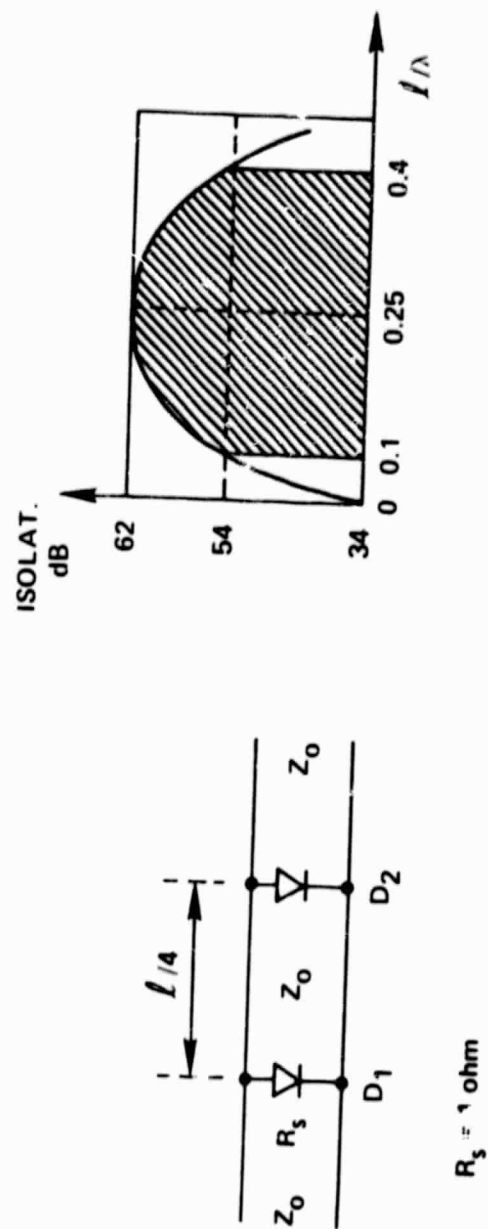
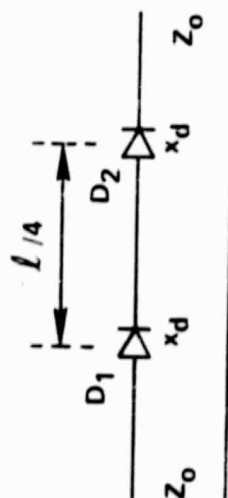
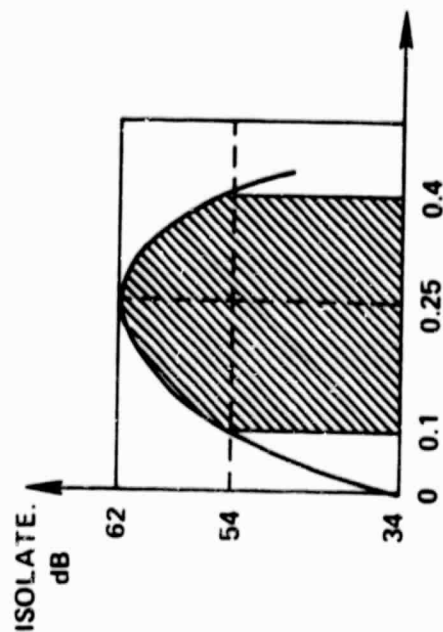


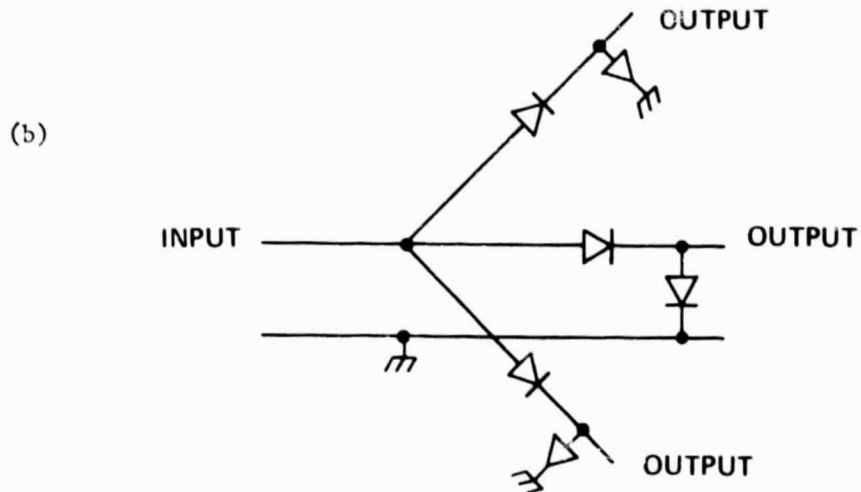
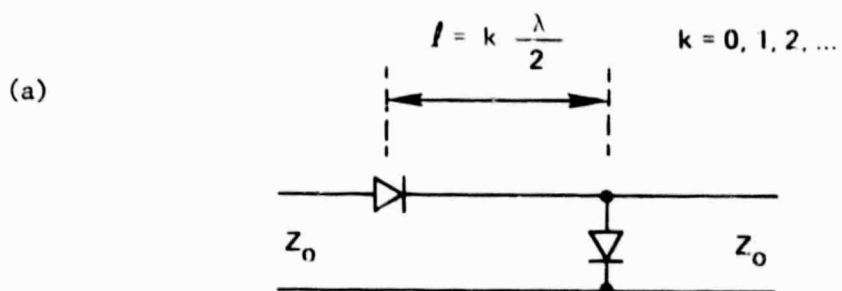
Figure 5-30. Shunt Diodes Quarter Wave Spaced





$Z_0 = 50 \text{ ohms}$   
 $x_d = 2500 \text{ ohms}$  ( $C_d = 0.0 \text{ pF}$ )  
 $f_0 = 3.2 \text{ GHz}$

Figure 5-31. Series Diodes Quarter Wave Spaced



1A-54,812

Figure 5-32. Series Shunt Switches

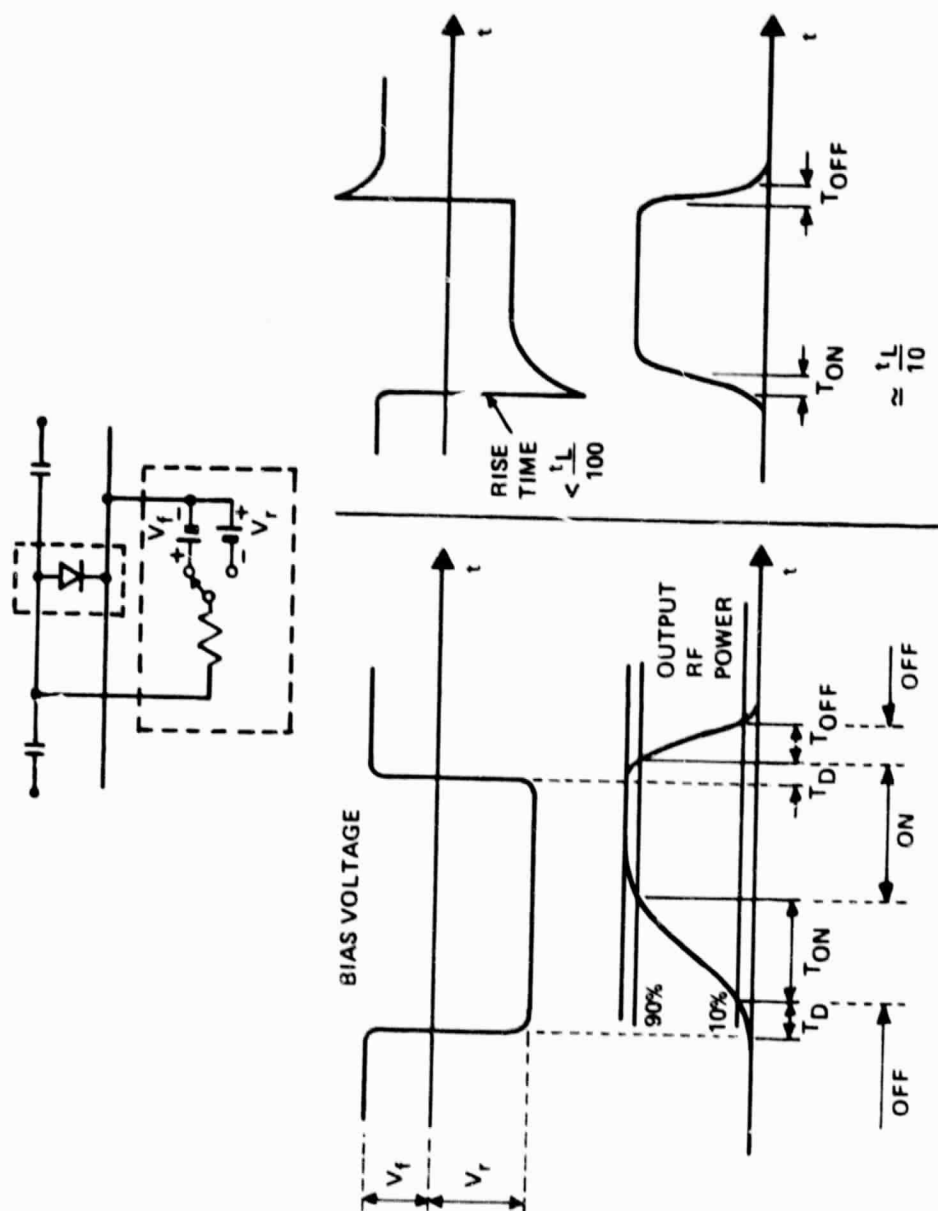


Figure 5-33. Shunt Diode Biasing and Switching

transmission.  $T_{ON}$  is also called reverse recovery time and is shown in figure 5-34 for two HP diodes: the 5082-3141, fast switching PIN diode and the 5082-3140, regular switching PIN diode.

As the above curves show, the switching time is related to the type of diode used; namely, it is proportional to the charge carrier life time  $t_L$ . The switching time is proportional to the forward current,  $I_C$  and is inversely proportional to the reverse voltage  $V_R$ . To reduce the switching time, the forward current has to be limited to the smallest possible value which is consistent with the required isolation (for shunt-diode switch) or insertion loss (for series-diode switch). Figure 5-34 also shows that a large reverse voltage reduces the switching time.  $V_R$  is limited in practice by the driver circuit, available power supplies, and the diode reverse breakdown voltage. From this point of view, a desirable value for  $V_R$  is -5 volts (consistent with a  $\pm 5$  V power supply).  $V_R$  is restricted to 0.8 volts in the series-shunt diode switch shown in figure 5-35 by clamping action of the ON-diode, which shunts the OFF-diode. The series-shunt switch has poor switching speed, unless the two diodes are d.c. decoupled and separately driven.

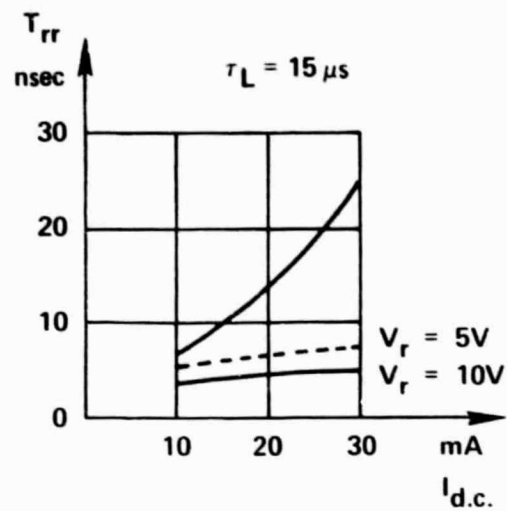
The bias voltage waveshape has to be modified to speed up the switching time, as shown in figure 5-33, by adding spike transients to the driving waveforms. The requirements of this waveshape are as follows:

1. Steep rise time: approximately  $t_L/100$ .
2. Duration of the transient: longer than the switching time.
3. Forward voltage spike: producing a forward current spike of magnitude which can be handled safely by the diode and the biasing driver.
4. Reverse voltage spike which does not exceed the reverse breakdown voltage of the diode (40 volts for fast-switching diodes and 100 volts for regular PIN diodes).

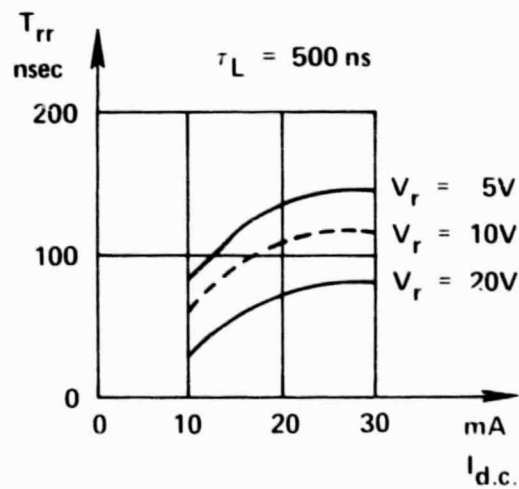
The resulting switching times may be as low as  $t_L/10$ .

The biasing circuit has to meet the following requirements:

1. It should have minimal effect on the insertion loss and the VSWR of the switch.



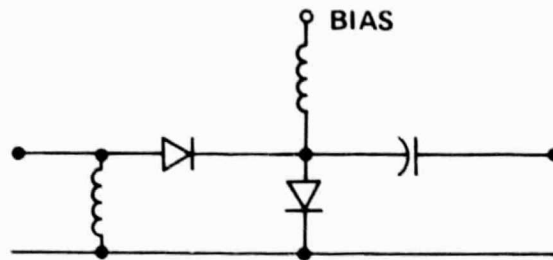
FAST SWITCHING DIODE, HP 5082-3141



SWITCHING DIODE, HP 5082-3140

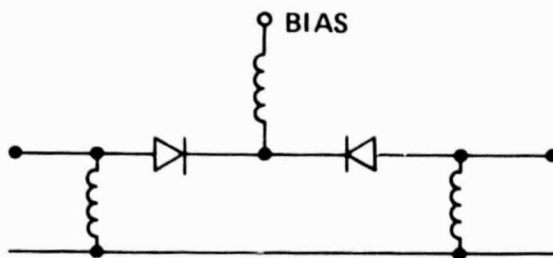
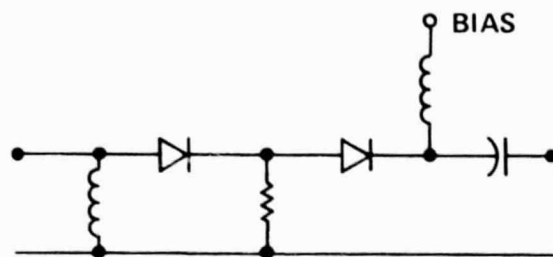
Figure 5-34. Reverse Recovery Time for Two Diodes

(a)



SERIES-SHUNT

(b)



SERIES-SERIES

1A-34,803

Figure 5-35. Clamping Susceptibility

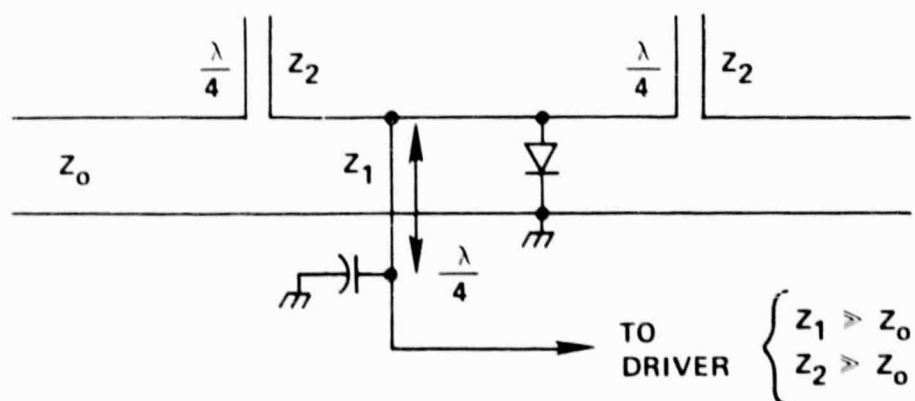
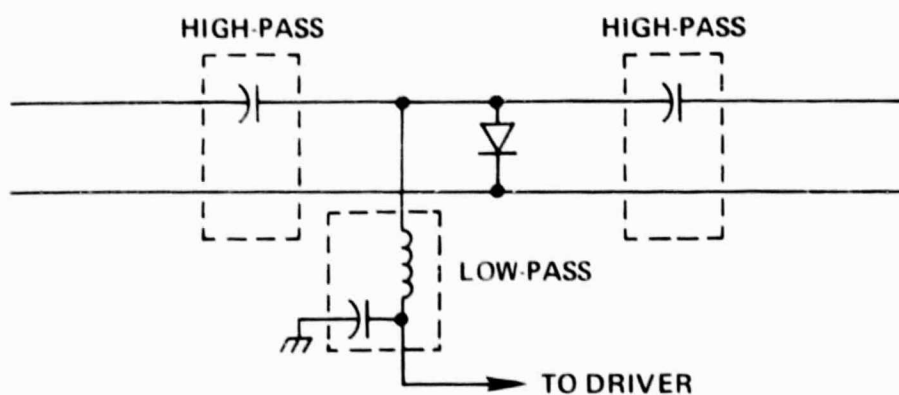
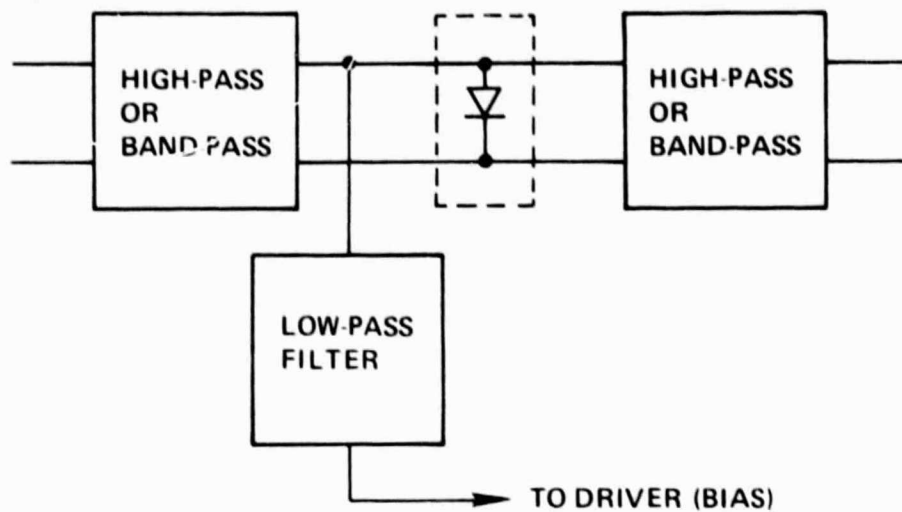
2. It should isolate the RF signal from the biasing circuitry. This is accomplished by using a low-pass filter between the RF path and the bias drivers. The low pass filter, however, should not appreciably slow the biasing waveshape.
3. The sharp bias spikes necessary for fast switching have to be prevented from reaching the signal output of the switch because they have microwave frequency components which contribute to in-band noise. This is accomplished by using high-pass or band-pass structures at the input and the output of the switch, as shown in figure 5-36.

The practical realization of the high-pass filter can be as simple as a series blocking capacitance, whose value is as small as possible, yet consistent with the lowest pass-band frequency of the switch. Instead of lumped elements, the blocking capacitor or the shunt inductor can be realized as  $\lambda/4$  transmission lines, also shown in figure 5-36.

Sufficient isolation can be obtained in the microwave region by using multiple-diode switches. While no particular biasing problems are encountered in the shunt-shunt diode switch (figure 5-30, with biasing shown in figure 5-36), the series-shunt diode switch (figure 5-35a) and the series-series diode switch (figure 5-35b) have specific biasing problems. In the series-shunt switch, one of the diodes is forward biased while the other is reverse biased (see figure 5-35a). The ON-diode clamps the OFF-diode: they appear in parallel for the biasing voltage, preventing a high reverse voltage from being developed across it, thus slowing switching. The two diodes have to be d.c. decoupled and biased separately to overcome this problem. In the series-series switch (see figure 5-35b for the biasing circuit) the faster of the two diodes will stop conducting sooner during the transition from ON-state to OFF-state blocking the discharge current of the slower diode. To prevent this, a high value resistor can be added or bias can be supplied at the mid point between the two diodes, as shown in figure 5-35b. The latter method is not applicable when more than two diodes comprise the series-series switch.

#### 5.2.5 Single Pole N Throw Switch

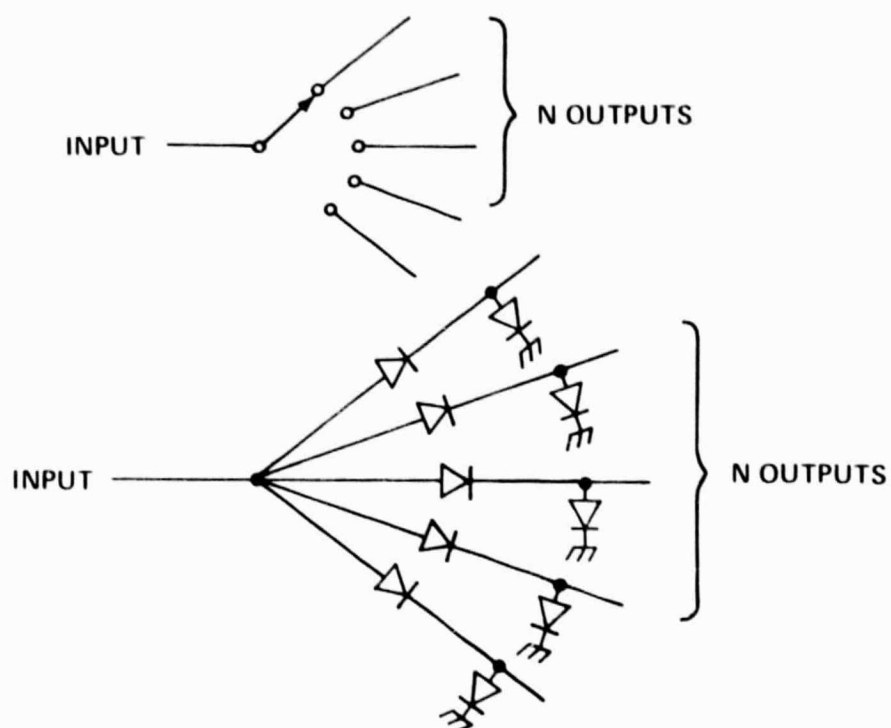
The practical realization of a single pole N throw switch is not very different from its symbolic representation shown in figure 5-37. Although N output arms of the switch may contain any number of series or parallel diodes, to increase isolation, the input to each arm is always a series diode which isolates the shunt diodes in the arm from the common input port.



1A-54,806

Figure 5-36. Filtering Against Spikes





1A-54,807

Figure 5-37. Single Pole N Throw Diode Switch

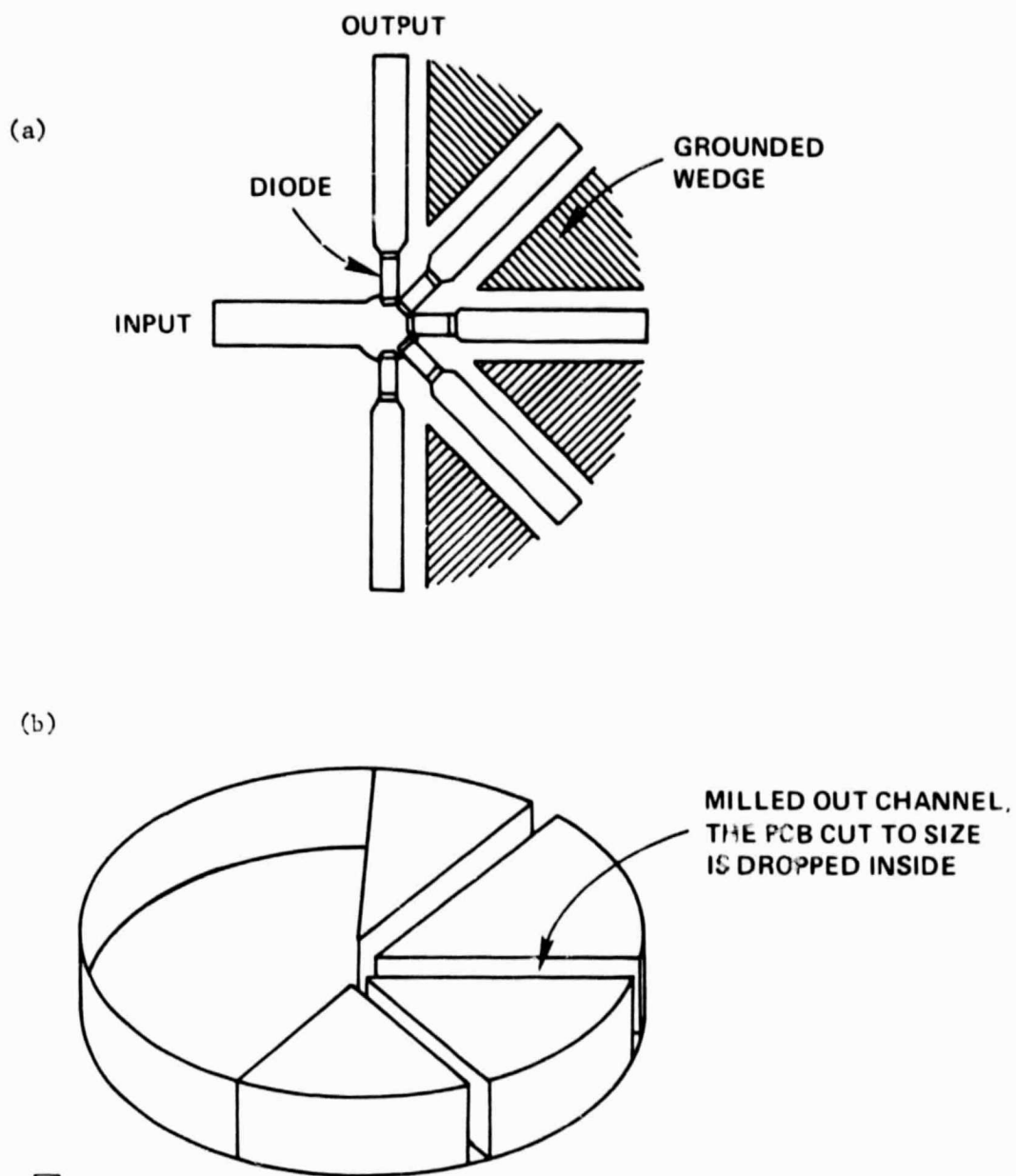
As shown in figure 5-37, the input series diodes are clustered together. The absence of transmission between the input junction and the  $(N - 1)$  disconnected lines accounts for the broadband performance of such a switch. The input line of characteristic impedance  $Z_0$  is terminated by an impedance  $Z_0$  because only a single output line is connected.

An important difference between the symbolic representation of the single pole  $N$  throw switch and the diode realization is that in the symbolic representation, the input can be connected to only one output at a time, while in the diode switch the input can be connected to several outputs simultaneously by appropriate biasing. As more than one output line is connected to the input line, the input impedance, looking into the switch, will decrease accordingly. To overcome this, an isolator may have to be used in the input line.

When broadcasting is done with the single pole  $N$  throw diode switch, the signal level at the output of the switch will not be constant but will vary due to power division and impedance mismatch effects, depending on how many outputs are connected simultaneously to the input. Since it is not known at the present time how many outputs may be connected simultaneously, the large signal level variations which may result preclude the broadcast mode by the  $1 \times N$  switch. The single pole  $N$  throw switch can only have simultaneous broadcast externally added to it.

The electrical performance (insertion loss, isolation, switching time) of a  $1 \times N$  switch is not very different from a single pole single throw switch. The isolation of the switch can be as high as 60 to 70 dB even in the GHz region by using more diodes in each output arm of the switch and by properly decoupling the output lines as they fan out from the common input junction. One way of decoupling the lines is to use grounded wedges between them (see figure 5-38a, and another way is real screening between the lines. This is done by making the whole switch from a solid piece of metal and milling out channels for each output line (see figure 5-38b). The printed circuit board (PCB) with appropriate pattern printed on it is cut to size and dropped between the channels. The diodes are bonded and then a cover is epoxied. Such designs are used by Omni-Spectra (model 2436-6826, etc.) and provide isolation in excess of 60 dB up to 18 GHz.

The insertion loss of a  $1 \times N$  switch increases monotonically with frequency. The reason is that the series-diode and shunt-diode switches which comprise each arm of the switch behave in the same manner, as shown previously in figures 5-26 and 5-28. The insertion loss depends on the isolation; a higher isolation will require more diodes, which in turn will increase the loss; a higher switching



1A-34,808

Figure 5-38. Grounded Wedge and Milled Decoupling

speed will require less forward biasing current which will increase the diode resistance and, with it, the insertion loss. Typically, an insertion loss of less than 1.5 dB can be expected from a switch with 60 dB isolation and 500 ns switching time for frequencies below 10 GHz. For faster operation, the insertion loss may increase to 2 dB.

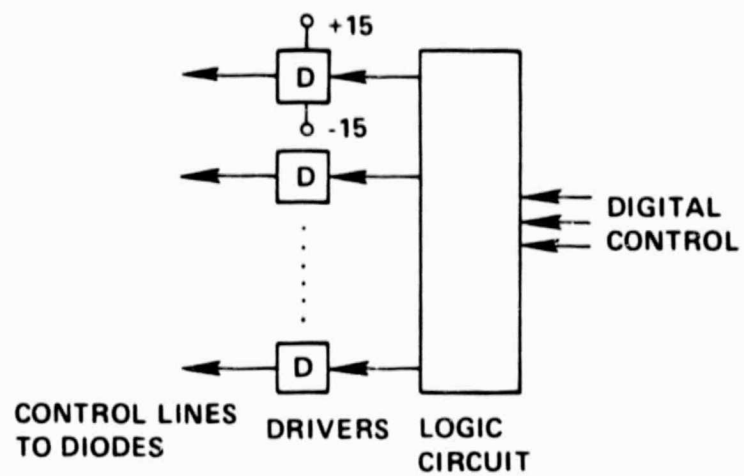
The input VSWR of a  $1 \times N$  switch depends on the number of outputs,  $N$ . The reverse bias capacitances of the  $(N - 1)$  open series diodes shunts the input port. For  $C_d = 0.05$  pF and  $N = 10$  ( $1 \times 10$  switch) the total capacity shunting the input is  $(10 - 1) \times 0.05 = 0.45$  pF. This will result in a VSWR of 1.88:1 at 4 GHz. Some inductive compensation (forming a low-pass filter) is possible, resulting in a 1.2:1 VSWR up to 6 GHz.

Practical limitations restrict the size of a single pole/ $N$  throw switch to  $N$  less than 10. For  $N$  higher than 10 the switch will have to be realized as a cascaded connection of simpler switches, as shown in figure 5-8 which depicts a single pole 100 throw switch, using  $1 \times 10$  switches as building blocks. In such a case, the total number of  $1 \times 10$  switches is equal to  $N + 1 = 11$ . The insertion loss is twice the insertion loss of a single  $1 \times 10$  switch, about 4 to 5 dB for each patch in the  $1 \times 100$  switch. Each (simpler) switch has to provide an isolation equal to the isolation required from the composite switch.

The number of bias control lines in a  $1 \times N$  simple switch (figure 5-22) equals the number of outputs  $N$ . In a cascaded switch (figure 5-8) the number of bias control lines is  $N + N^2$  where  $N$  is the number of outputs of the building block. Each control line will supply each arm of the switch with +5 to +15 volts or -5 to -15 volts in order to turn it ON or OFF. Logic will be incorporated in the circuit in order to ensure that when one arm of the switch is ON, the remaining arms are OFF. The most logical arrangement is to use as many drivers as there are control lines and to precede the drivers with logic circuits, as shown in figure 5-39.

#### 5.2.6 Driving Considerations

The characteristics of a PIN diode would be optimized for use in a cross bar switch. The reverse bias capacitance and the forward bias resistance will be adjusted as a function of driving current and switching time within the possibilities of the state of PIN diode art to secure the best combination of isolation, insertion loss, switching time, and drive power consumption. A special order may be placed for several thousands of custom made diodes. With this sort of control, what power consumption can be expected of the



1A-54,803

Figure 5-39. Sequence Digital Control, Decode, and Drivers

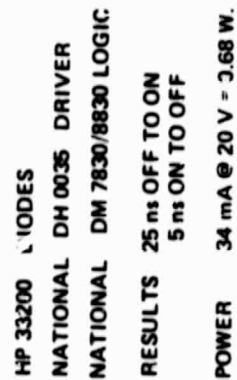
diodes? An answer to this can be estimated by considering four different drivers used in current practice.

The first is shown in figure 5-40, and is comprised of a National DM 7830/8830 logic receiver, and a National DH 0035 driver, used to control two HP 33200 shunt diodes. These are forward biased when the switch is open. It takes the driver 25 ns to sweep the charge out of the I region with the aid of recombination to close the switch. Reinserting the charge to open the switch is faster, only 5 ns. The power consumed under forward bias is 34 mA at 20 V or 684 mW. This is a high number for projecting up to 10,000 cross points which would then use 6.84 kW and 340 A. The driver is acceptable for a single switch using two diodes, but obviously not efficient enough for a large switch.

This performance should be compared with the Omni Spectra tentative specification of figure 5-11 which estimates 150 A and only 750 W for a 100 x 100 cross bar. Each switch requires 15 mA.

Another switch recently developed at TRW [Salcedo, 1979] uses five transfer switches to perform a 4 x 4 crossbar function. Each transfer switch is equivalent to four series-shunt switches and requires two drivers each having dual inverted outputs. Thus, there is an output for each series-shunt switch. These are fed approximately 20 mW (10 mA at 2 V) so a transfer switch requires about 80 mW. The raw input to each of two drivers per transfer switch is 130 mW for a total of 260 mW. The ratio of raw power to diode power is thus 3.25 to 1. If these results are applied to a 10,000 cross point switch, diode power would be 200 W and the raw power 650 W, about the same as the switch vendor estimate. Isolation in a transfer switch is 56 dB.

A final low powered driver was reported by Glance of Bell Labs [Glance, 1979]. It is designed to operate a 4-bit (coupler-type) phase shifter using two series diodes per bit. HP beam leaded diodes (HPND-4001) are used with a capacitance of 0.07 pF at -30 V and series resistance of 25 ohms at 10 mA. The diodes of a bit use 5 mA of forward current and can be switched on or off in 1.0 ns. The d.c. power consumption of a cell for both diode and driver is 8.75 mW, a low figure. If this could be scaled to a 10,000 cross point switch, the total raw power required would be only 87.5 W. This is a factor of seven reduction over the TRW and Omni Spectra results, but should be tempered by the knowledge that the Bell Labs phase shifter uses series switches only so that no current is drawn by a back biased cell. Circuit details are proprietary and have not yet been released for publication.



**Figure 5-40. Driver for Shunt Switches HP**

The raw power required to operate the drivers of a 100 x 100 switch using PIN diodes is about 500 W. Details of the efficient driver of the Bell Laboratories have not been made public, but are expected to be published soon by N. Amitay.

#### 5.2.7 PIN Diode Cross Bar Conclusions

A PIN diode cross bar can be designed to meet the tentative electrical specifications for the 100 x 100 switch. Mechanical problems, such as packaging and heat dissipation need study. Power consumption for drivers and diodes is expected to be about 500 W. The PIN diode switch is the most mature electronic switching technology for microwaves but may be bulkier and require more power than newer switching technologies.

#### 5.3 FERRITE TECHNOLOGY

Ferrite devices are used in switches, isolators, isolation phase shifters, and a component unique to them, circulators. A circulator is possible with ferrites because they are non reciprocal materials. Circulators permit the construction of a ferrite switch. Figure 5-41 shows the outward behavior of a four port circulator and helps to make clear some of its uses. The unit in the figure connects signals among its ports, numbered from 1 to 4, in a circular fashion. A signal entering port 1 is directed to port 2, one entering port 2 is conducted to port 3, and so on till a signal entering port 4 would be returned to port 1. Such behavior clearly violates reciprocity since a signal from port 2 is isolated from port 1 by about 20 dB. The circulator in figure 5-41 is used to connect both a high powered transmitter and a low noise amplifier to a single antenna. It helps to match the transmitter and to protect the low noise amplifier, since the transmitter power will all leave via the antenna if that unit is well matched. If it is not, the power reflected by the antenna will not return to disturb the transmitter. Instead, it will proceed to port 3, be rejected by the preselector filter, and proceed to the load at port 4 where it is absorbed by a termination. Signals picked up by the antenna in the receive band enter port 2, are directed to port 3, are passed by the preselector, and enter the low noise amplifier.

All the foregoing is possible because the circulation of signals among the ports is clockwise. This sense of circulation is determined by the direction of a magnetic field existing in the ferrite material in the circulator. If that field were reversed in direction, the sense of circulation would be counter clockwise. The results would be inconvenient for the transceiver of figure 5-41,



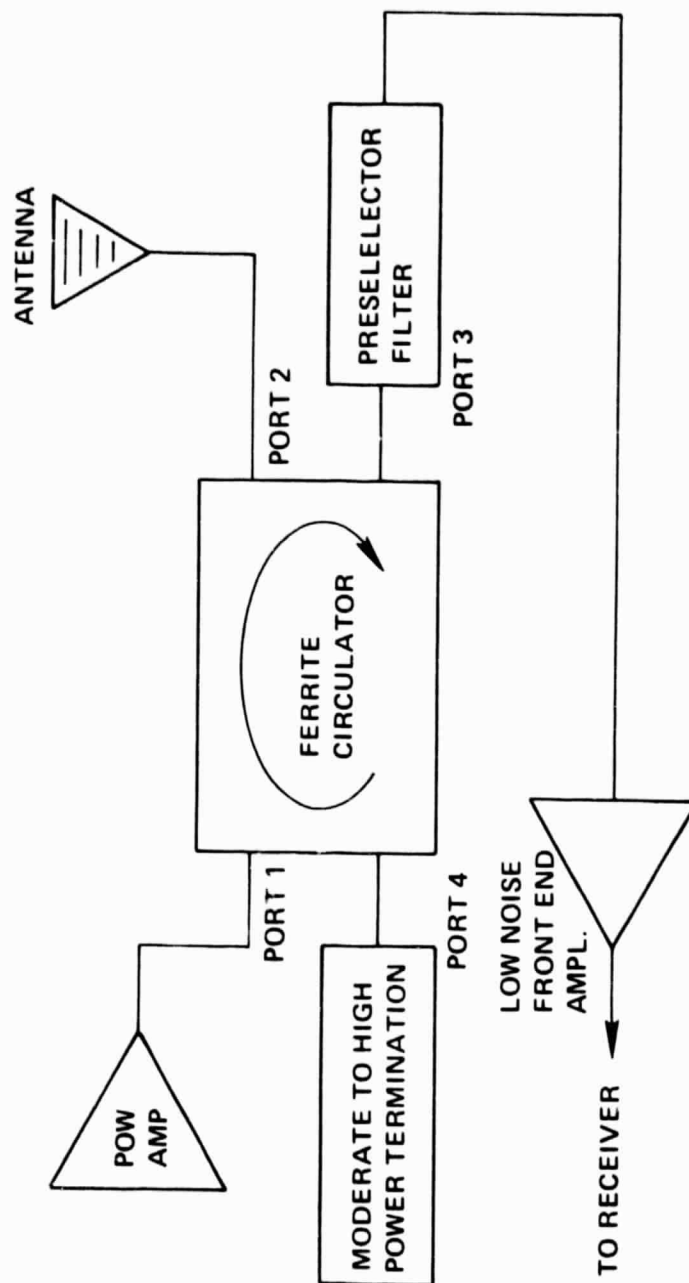


Figure 5-41. Circulation Principles

but not for a circuit intended for switching. Thus, the basis for ferrite switching is to reverse the magnetic field biasing the ferrite material and change the sense of circulation.

A three port Y-junction circulator small enough for microstrip or stripline use has been developed. Figure 5-42 shows the geometry of such a circulator. The junction takes the form of a Y with three input lines separated by 120 degrees. Circulation can be achieved in either clockwise or counter-clockwise directions by reversing the magnetic field normal to the plane of the figure. Two methods are used to produce the normal magnetic field which is taken, by convention, to be in the Z direction. The first is application of an external magnetic field by means of an electromagnet or by addition of a small high strength permanent magnet to the structure outside the ground plane. The second is providing a closed d.c. magnetic circuit in the ferrite material itself and utilizing the remanant magnetic field of the hysteresis loop of the material. In the second method, high reluctance gaps are avoided and a lower driving magneto motive force is required to set up the fields. This is supplied by enclosing a loop of wire in the ferrite itself. The wire geometry must be carefully controlled so that only a simple field parallel to the Z axis is induced. The low reluctance and small size enables the field to be set up or reversed in a time as short as a microsecond, a speed which is suitable to radar phase shifter applications.

#### 5.3.1 Microstrip and Stripline Circulators

A fairly good theory exists for the operation of printed microwave circulators. It was discovered and described by [Bosma, 1964], amplified by [Fay and Comstock, 1965], and adapted for microstrip by [Rosenbaum, 1974]. A detailed exposition of the theory is not given here. Factors important for switching purposes are recounted in the form of an outline. In short, Maxwell's equations are solved for two cylindrical ferrite discs encased in a stripline medium as shown in figure 5-43. Since the discs are symmetrical, the modes need only be found in one disc and, if fringing can be neglected, the solution can be applied to microstrip as well. The analysis finds two contra rotating modes producing standing wave fields which couple adjacent ports in cyclical clockwise or counter-clockwise order, depending on magnetic field direction.

A design procedure based on the Bosma theory has been determined and it is outlined in the references cited. Certain printed circuit designs based on this theory have been reported and these have determined material, disc radius, field, and conductor

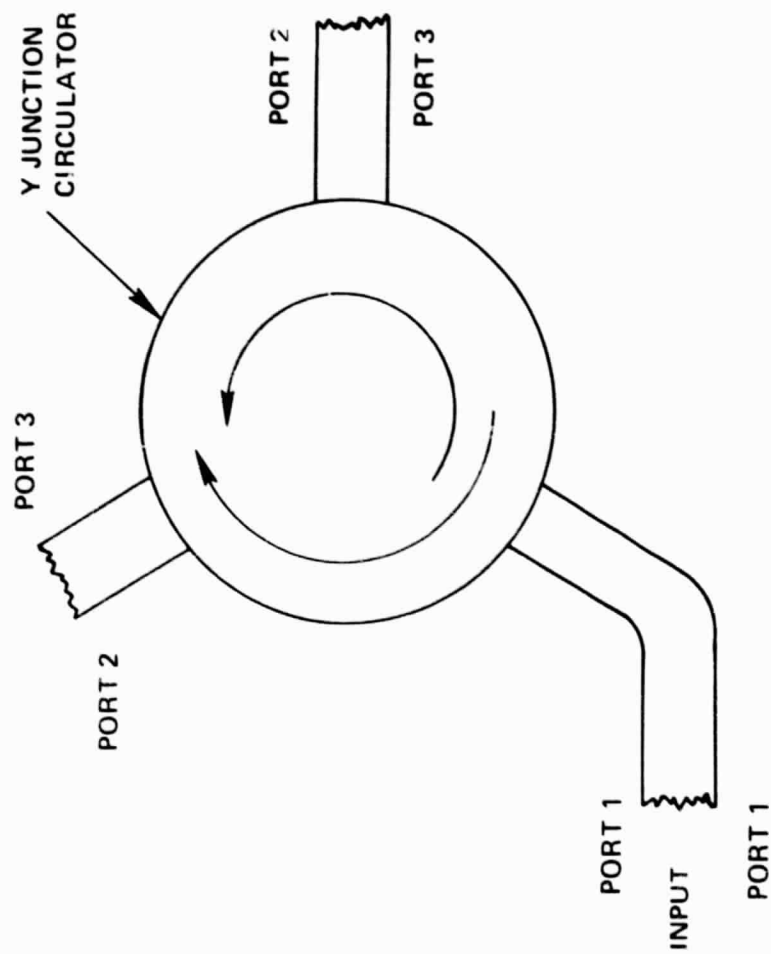


Figure 5-42. Y-Junction Microstrip Circulator

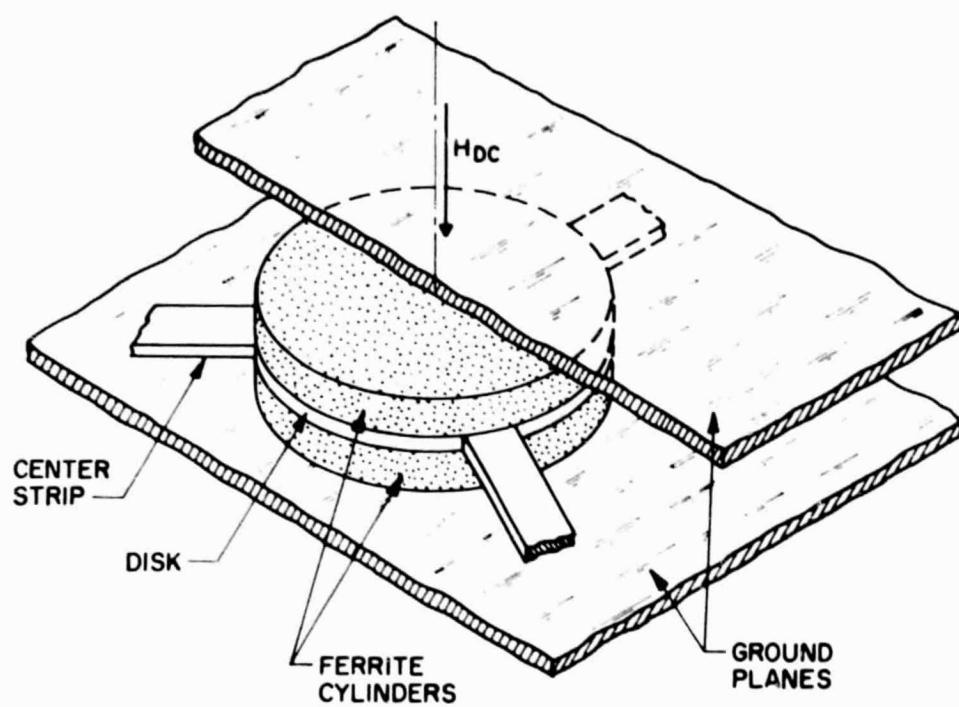


Figure 5-43. Stripline Y-Junction Circulator

width-subtended angle as suggested by figure 5-44. An important paper [Wu and Rosenbaum, 1974] shows how to get microstrip circulator bandwidths of 76 percent with an isolation of 15 dB and an insertion loss of 1.0 dB when operating at a center frequency of 9.0 GHz. The paper is an extension of the Bosma, Fay and Cornstock theory and discloses the fact that wide transmission line coupling to the ferrite circulator (i.e.,  $\psi = 0.5$  radians) is a significant design parameter. It is unfortunate that this paper was published after a number of latching stripline and microstrip circulators were developed for various applications, because the best techniques have not all been applied to one design. What can be accomplished with a microstrip latching ferrite circulator switch is a matter of judgment based on a review of reported stripline/microstrip circulators.

### 5.3.2 Printed Circuit Circulators

A latching ferrite circulator switch in stripline [Betts, et al., 1966] was temperature stable and could handle a high peak power. It operated at a center frequency of 2.9 GHz and had a 26 dB isolation bandwidth of 240 MHz (8.9 percent). The magnetic field was the remanent field of the ferrite (240 G) after magnetic pulsing. Quarter wave matching into the ferrite disc was employed. The diameter of the ferrite disc was 0.705 inch. The device had an insertion loss of about 0.35 dB, and could be switched in 10  $\mu$ s. The switching energy required was 450  $\mu$ J. An external magnetic circuit was brought through the ground planes. While conductivities were maintained for the RF field, a slit was made in the ground plane to eliminate the shorted turn effect.

A microstrip circulator built entirely on ferrite in two pieces with an embedded wire for magnetic switching drive was reported by [Hershenov, 1967]. The d.c. magnetic field was almost completely contained in the ferrite, and the disc was defined by the central conductor metalization. The circulator was self latching and had a 10 dB isolation bandwidth of 350 MHz, probably due to incomplete understanding of the perturbing effects of the driving wire and a non ferrite dielectric insert which held the wire. A combination of this work with that of Wu and Rosenbaum, whose circulator was also built completely on ferrite, was needed.

To improve on the Hershenov approach, [Siekanowicz and Schilling, 1968] completed the magnetic circuit with a hollow ferrite cylinder outside of the central ferrite disc. A non magnetic dielectric cylinder separated the two ferrites and supported the driver wire used for reversing the magnetic field. They achieved a

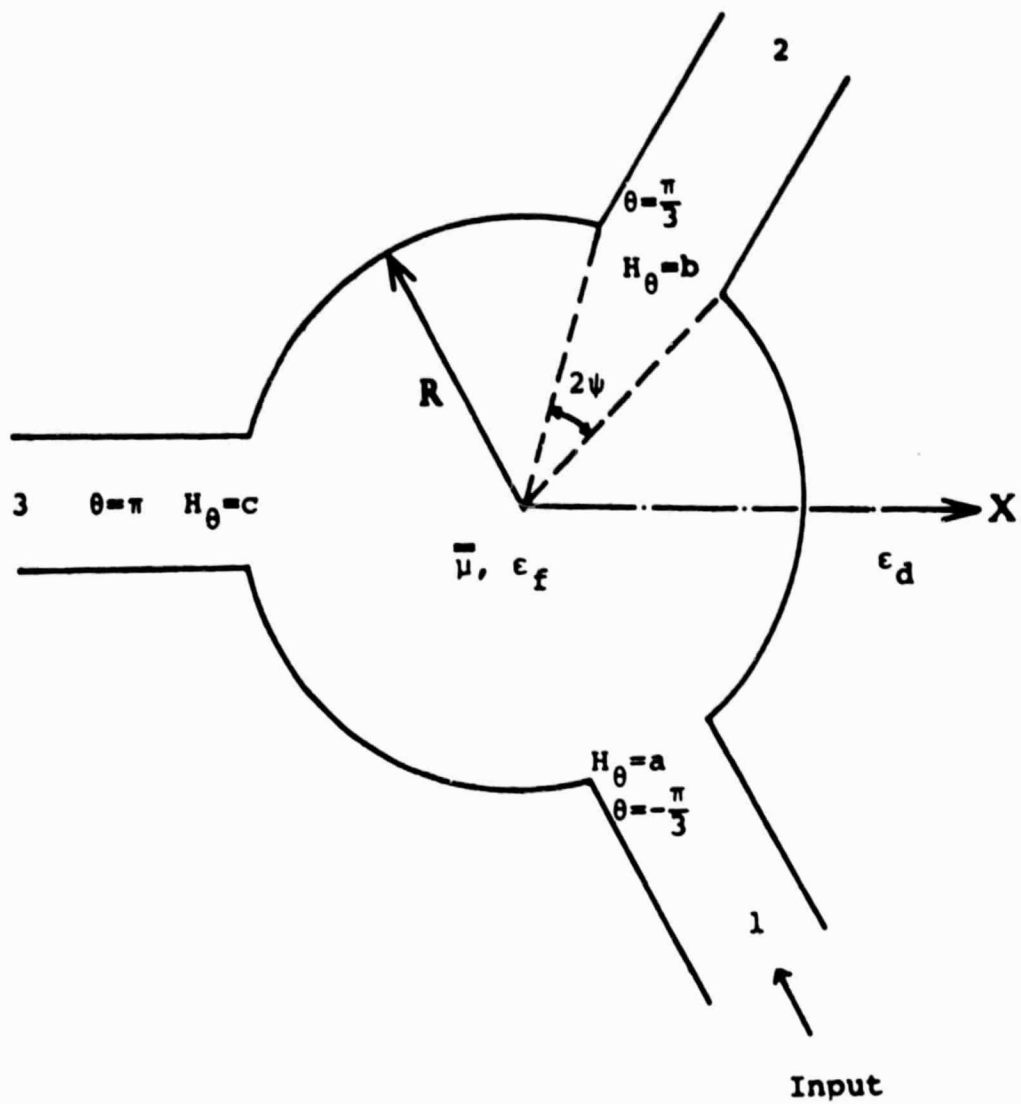


Figure 5-44. Geometry of Center Conductor Pattern of Y Circulator

20 dB isolation bandwidth of 4.1 percent centered at 7.5 GHz. The switching time and energy for this circulator were 0.2  $\mu$ s and 30  $\mu$ J.

[Ayter and Ayali, 1973] used the Wu and Rosenbaum technique to build a stripline circulator which operated over an octave band with a 15 dB isolation and 1 dB insertion loss. Their objective was to derive useful wideband design curves.

Most recently, [Helszajn, et al., 1979] analyzed stripline triangular ferrite circulators and pointed out that these circulators can theoretically have a bandwidth three times that of a circulator employing a disc shaped ferrite.

### 5.3.3 Ferrite Switch Design

What kind of switch could be designed with printed circuit ferrite circulators? Bandwidth would not be a problem because of the newer discoveries of Wu and Rosenbaum. Losses per circulator run 1.0 to 1.5 dB on the average. If proper architectures are used and high enough signal levels are applied at the input to the switch, these losses are acceptable. Further, when using ferrites, the switch frequency need not be translated down as far as 3.0 GHz from the 30 MHz uplink frequency, since printed ferrite circulators will perform well in the 7 to 9 GHz band. Failure rates for latching ferrite devices are expected to be low since mechanical damage is the main cause of failure.

Disadvantages of ferrite switches are as follows. First, the isolation for one circulator is low, 15 dB over the band. A switch cross point requires at least three circulators, as shown in figure 5-45, to provide connections for paths guiding input signals across the matrix and those guiding signals upward to connect to downlink beams. The isolation in this case would only be 30 dB which is barely adequate to produce an acceptable signal to cross talk ratio. The cross point shown in figure 5-45 is truly planar, does not require even a second layer, but is too large, due to matching. Every circulator must be matched if it is to produce isolation, otherwise the output signal is reflected and circulates to the isolated port. Transformers are built into all broadband circulators to provide the match. Broadband circulators reported are about 1 inch in diameter. Thus, the cross point in figure 5-45 has an area of at least three square inches which would drive the volume of a 100 x 100 switch up to at least 15,000 cubic inches, three times as large as with semi-conductors.

A further disadvantage of the ferrite switch is switching speed and drive power. Printed ferrite circulators are reported which

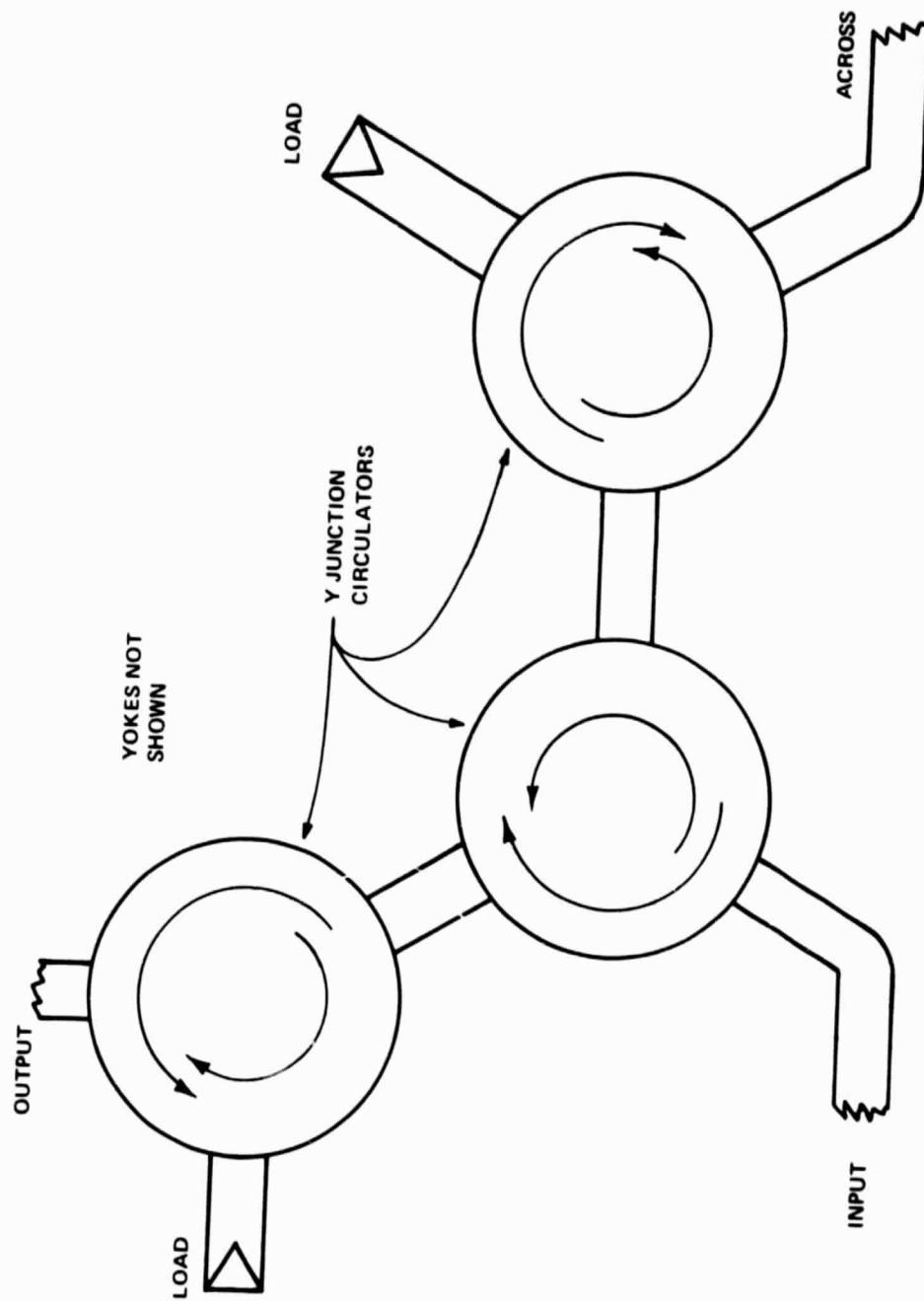


Figure 5-45. Y-Junction Ferrite Cross Point



switch in 0.5 to 1.0  $\mu$ s. This is too slow by an order of magnitude. Switching speed is reduced because of the necessity of building up a driving magnetic field in a volume of 1/4 to 1/2 cubic inch with a strength of 1000 Gauss or more through inductive circuits, and through metallic shields which carry induced eddy currents. The field must also propagate through the ferrite. The average energy per switching operation is reported as 10  $\mu$ J. If a cross point consists of at least three circulators, as shown in figure 5-45, and if 200 cross points changed state on the average of once every 5  $\mu$ s, then the power, P, required to operate the switch would be

$$P = 10 \times 10^{-6} \times 3 \times 2 \times 10^2 / (5 \times 10^{-6}) \text{ W}$$

$$P = 1200 \text{ W,}$$

which is a very high power consumption. The driving current would be of the order of 1 A per circulator; thus, the minimum current needed for a change of switch state would be about 200 A delivered at a low average voltage. This requires a highly complex high current power supply to power the drivers for the circulators.

Ferrites have a minimum density around 5 grams per cubic centimeter. An individual circulator might have as much as 1  $\text{cm}^3$  of ferrite. Thus the minimum weight of ferrite alone in the switch would be 330 pounds. This does not include the rest of the switch such as connectors, transmission lines, structures, drivers, and logic. Ferrite devices suffer from intermodulation products which experience shows are 30 to 40 dB above those due to metal transmission structures.

#### 5.3.4 Ferrite Device Conclusions

Ferrite technology has too slow a switching time (0.5 to 1.0  $\mu$ s), and requires too much switching energy (1.2 kW). The cross points require multiple circulators because isolation is low and thus total switch volume rises to 15,000 cubic inches, at least three times larger than semi-conductors. The weight of the ferrite pellets in the circulators alone would total 330 pounds. The high current pulses totaling at least 200 A will generate EMI, and ferrite intermod performance is poor. These disadvantages do not make up for broadband operation (one octave), single layer layout, and high reliability. The conclusion is that ferrites cannot perform well enough for use in a switch.

## 5.4 OPTICAL COMMUNICATIONS TECHNOLOGY

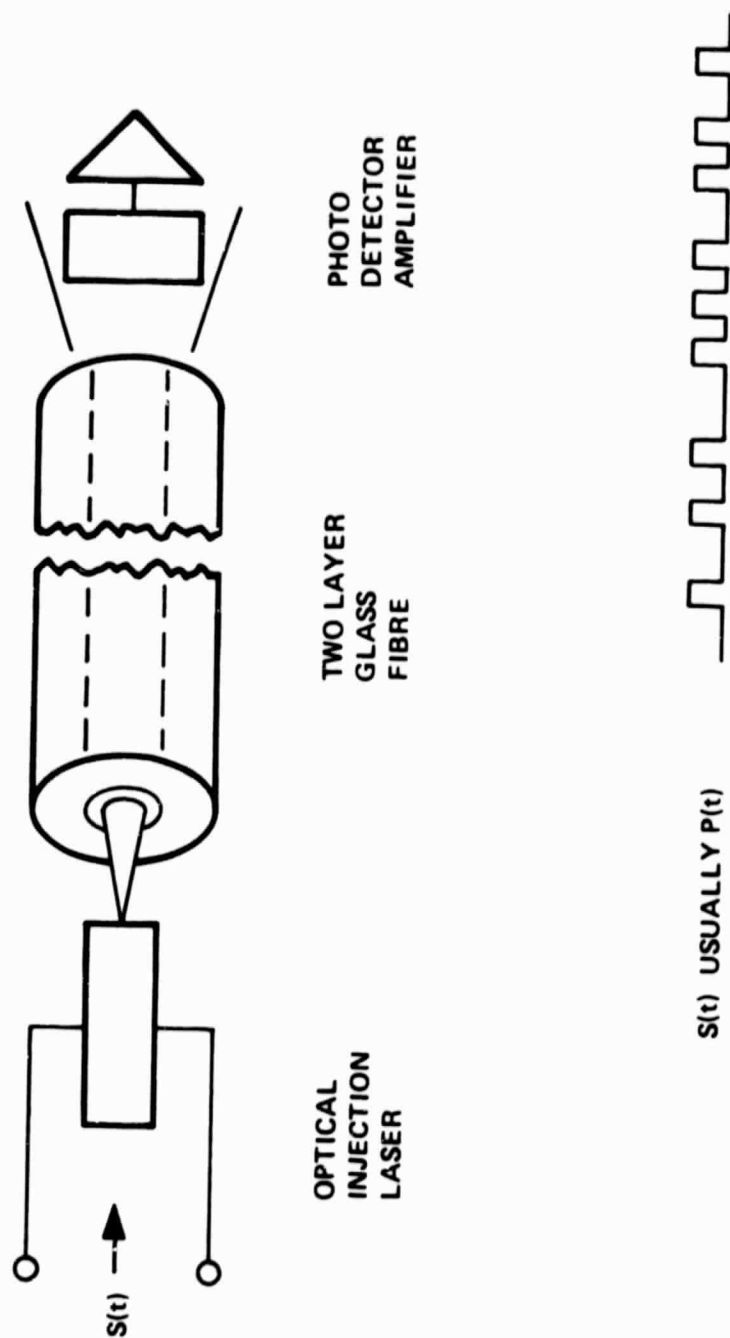
The essentials of optical communications technology are shown in figure 5-46. If optical switching is employed, the microwave signals will have to be translated to baseband so they can be used to amplitude modulate an optical carrier generated by an injection laser. This process is shown at the left of the figure. The injection laser produces a linear amplitude modulated optical signal. Available bandwidths of the modulator are about 100 MHz, but experiments are currently in progress on modulator bandwidths slightly in excess of 1 GHz [Bosch, 1979]. If a 2.5 GHz bandwidth were achieved, the uplink band centered around 30 GHz would then be translated downward to 1.250 GHz. The laser costs several thousand dollars and weighs about 5 pounds with its power conditioning and modulating circuits. One laser and frequency translator would be required for each of 100 inputs.

### 5.4.1 Optical Switching

In order to use an optical switch, it will be necessary to transfer from the microwave to the optical medium and to employ all the additional transducing equipment required. Thus, 100 injection lasers, input cables, output cables, photodetection systems, the necessary mounting, support, protection, conditioned power, and signal interfacing equipment would all be charged to the switch, which would begin and end at the interface equipment. Whether this added equipment plus that of the switch itself will total to a cost, weight, and volume which is competitive with other switching techniques cannot be determined at this time. Circuit integration and miniaturization may incorporate many of these functions on small chips at some date in the future.

The laser generated light beam exits in the form of a cone which impinges upon the end of a glass fiber transmission line. The line consists of a central core of high quality optical glass, whose attenuation is minimum at a  $0.9 \mu$  wave length, and an outer cylinder coat of stepped or graded lower index of refraction. Not all the light emitted by the laser is captured by the fiber because there are reflections and misalignment of the cone and the core. The light that enters the core propagates with a low loss of a few decibels per kilometer. It is trapped inside the core by total internal reflection at the interface between the core and its outer coating.

There is a range of angles of incidence at the core-coat interface; each angle corresponds to a different mode of propagation in the core with different phase and group velocities. The



BUT FM IS REPORTED WITH A BANDWIDTH OF 300 MHz FOR 1 km

Figure 5-46. Optical Switch Principles

different modes are excited by the variation in angle of the rays of light in the cone emitted by the laser. Consequently, the laser signal suffers phase distortion which increases as a function of distance traveled in the fiber. Thus system bandwidths are limited and are rated in megahertz per kilometer. Systems supporting as much as 100 MHz over distances of 8 km have been reported [Kane, 1979]. Much shorter distances would be required for an optical switch, so the phase distortion would not be as severe. Glass fibers which support higher order modes have had cores with diameters of 50  $\mu$  and covers with diameters of 125  $\mu$ . Newer fibers are being developed with smaller diameter cores to reduce moding and thus distortion.

The receiving end of an optical communications system consists of a broadband photo-diode detector and amplifier which transduces the signal back to the electronic circuits, which then make use of it. An optical switch would be located somewhere at the middle of the cable and would be only one of many cables entering the switch and leaving it.

The most appropriate form of cross bar switching at optical frequencies is described in a report of research done for Rome Air Development Center [Soref, 1979]. The switch utilizes the Pockels linear electro-optic effect which makes it possible to create waveguiding and switching structures within a single crystal plate of the ferro-electric material,  $\text{LiTaO}_3$ . Metal patterns form electrodes which apply voltages in the Z direction, perpendicular to the plate and which produce the changes in the index of refraction used for guiding and switching. Switching is accomplished by using total internal reflection beyond the critical angle at a voltage induced discontinuity in index of refraction. Almost uniquely among commercial electro-optic materials,  $\text{LiTaO}_3$  is available with a combination of the following desirable properties: large sizes, freedom from domain-striae-imperfections, high resistivity, high purity, stability, and polishability. In small quantities, a 6.0 x 1.0 x 0.008 cm plate costs \$250.

The plates are Z-cut with light propagating in the Y direction. A 75  $\mu$  thickness is chosen to match the core diameter, 60 to 85  $\mu$ , of the multi-mode fibers end coupled to the plate. The crystal is homogeneous in the absence of voltage apart from a temporary passive effect stemming from metal overlays, top and bottom of the plate, to which voltages are applied. With the voltages off, light in the crystal is free to spread out in the X - Y plane, though it is strongly trapped in the Z direction due to total internal reflection.

The applied E fields produce the desired X - Y guiding. The fields are mainly along Z and are fairly uniform. The crystal's refractive index in both the X and Z directions, perpendicular to the direction of propagation, is perturbed according to the relationships

$$\Delta n_x = \frac{1}{2} n_x^3 r_{13} E_z$$

$$\Delta n_z = \frac{1}{2} n_z^3 r_{33} E_z$$

where the electro-optic r coefficients for LiTaO<sub>3</sub> at  $\lambda = 6328 \text{ \AA}$  are

$$r_{33} = 30.3 \times 10^{-12} \text{ m/V}$$

$$r_{13} = 7.0 \times 10^{-12} \text{ m/V}$$

Figure 5-47 shows the field dependence of the change in index of refraction  $\Delta n_x$  and  $\Delta n_z$ . Opposed top and bottom strip electrodes permit setting up walls or barriers in the crystal that are rectangular volumes, parallel to Z, at which the spatial distribution of refractive index decreases rapidly in step like fashion.

Light will bounce off such barriers over a range of grazing incidence angles due to the phenomenon of total internal multimode reflection. One can make mirror like objects or optical channels from such parallel barriers. Both TE and TM light can be confined by total internal reflection at bounce angles that range from zero to the critical angles

$$\theta_c(\text{TE}) = (2\Delta n_x / n_x)^{1/2}$$

$$\theta_c(\text{TM}) = (2\Delta n_z / n_z)^{1/2}$$

The field dependence of both critical angles is presented in figure 5-48. The TM light (polarized parallel to the d.c. electric field  $E_z$  is more strongly confined, although the square root dependence of the critical angle reduces the difference so that channeling for TE light (polarized perpendicular to the d.c. electric field) is 48 percent as effective. While switching potentials are time varying, a steady bias voltage is required to induce and maintain the network of waveguides in the crystal.

Various losses enter into the operation of a cross point matrix. These include reflections at the fiber/crystal interface, a lower percentage of TE light guided, crystal thickness and fiber core not the same size, some spreading of field outside the electrically created waveguides, cross point loss, and failure to

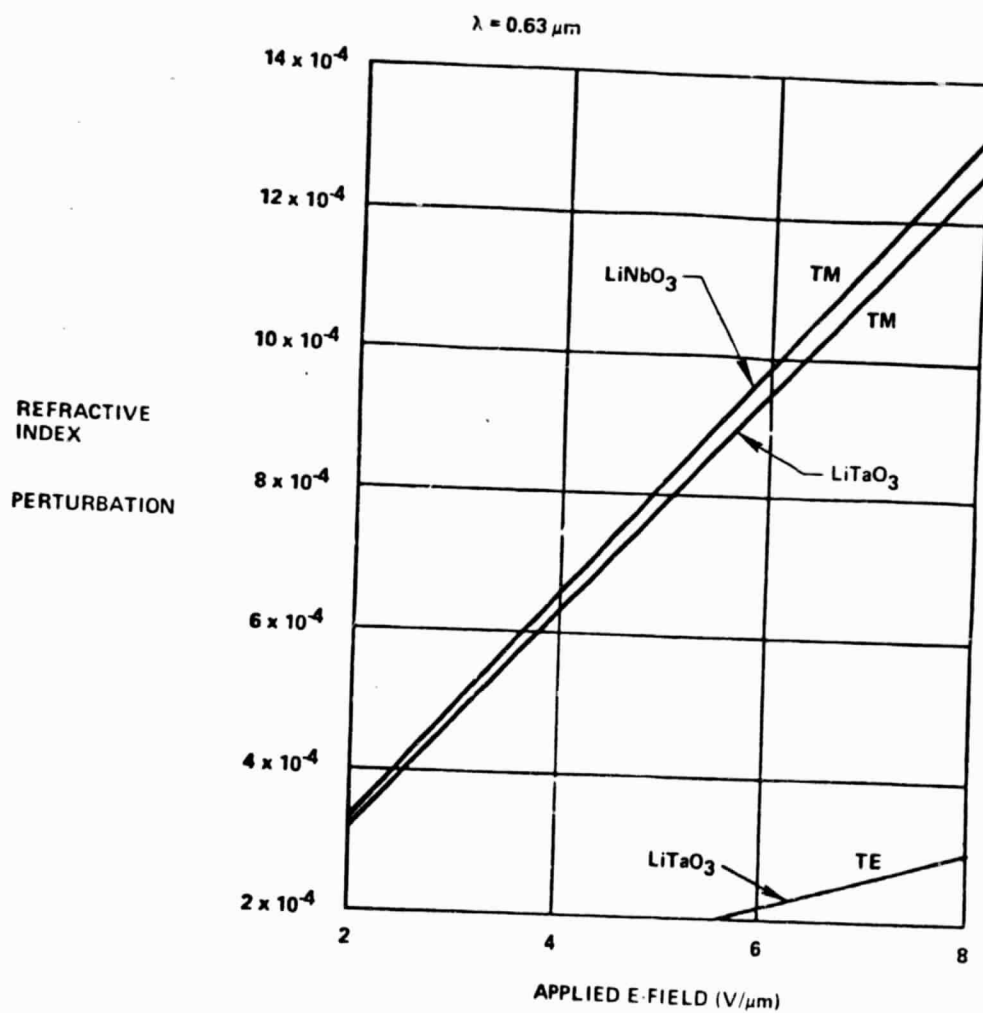


Figure 5-47. Index Perturbation LiTaO<sub>3</sub> as a Function of Voltage

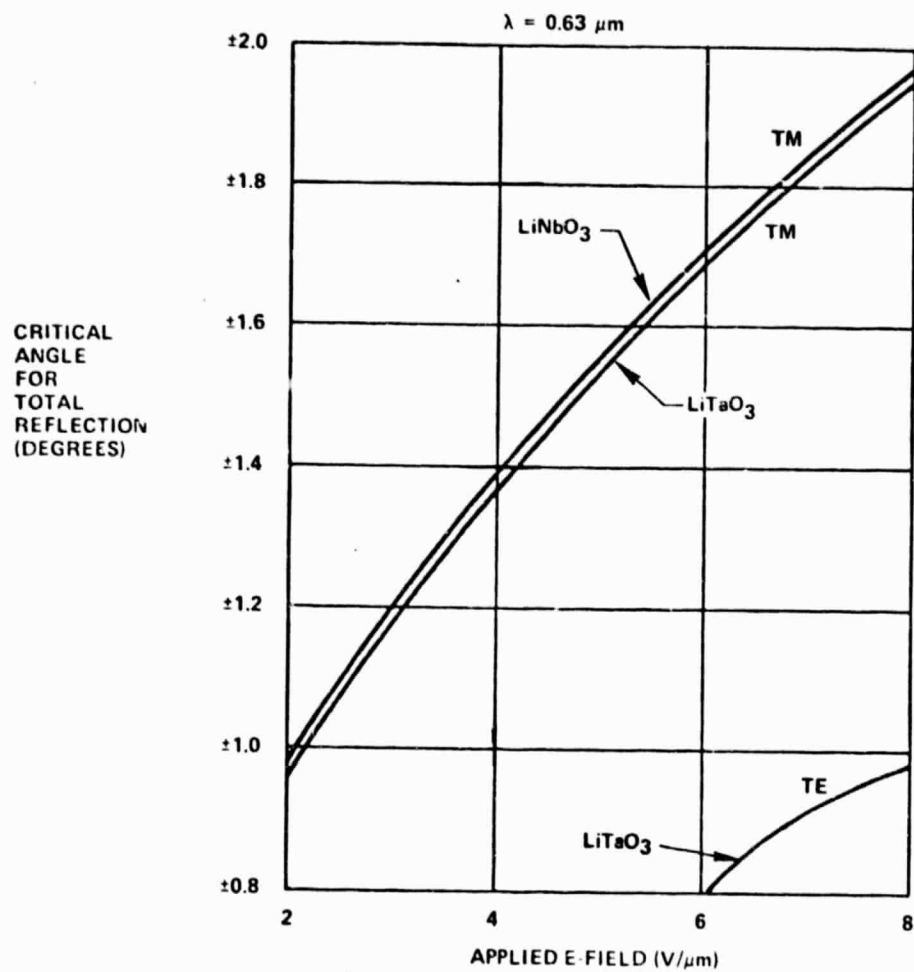


Figure 5-48. Field Dependence of Critical Angles



reflect all the power at the selected cross point. As a result, the estimated losses for a 3 x 3 cross bar matrix are 15.7 dB for TE and 11.7 dB for TM light.

Using the principles discussed, it is possible to develop a cross point 1 cm, long as illustrated in figure 5-49. Light entering the leftmost channel in the foreground will, if the deflection electrode is OFF, pass right through the cross point and continue in the channel, emerging in the electrode OFF direction shown in the figure. If the deflection electrode is ON, then the light is reflected into the upper output channel due to total internal reflection. It emerges in the electrode ON direction in the figure.

The significant characteristic of the optical switch is the employment of total internal reflection beyond the critical angle. This phenomenon should be briefly reviewed. When a light ray goes from an optically more dense medium to an optically less dense medium it bends away from the normal to the interface as shown in figure 5-50a. As  $\theta_n$  becomes larger,  $\theta'_n$  becomes smaller, and  $\theta_L$  reaches a stage (since it must be larger) where it is parallel to the interface. Beyond this point, the ray can not leave the optically more dense medium and so it reflects off the interface. These latter two stages are shown in figure 5-50b. The optical density is the index of refraction of the medium. It is equal to the square root of the relative dielectric constant.

The angle of incidence is measured relative to the interface normal. The grazing angle is its complement and as the incidence angle  $\theta_n$  gets larger and exceeds the critical angle, the grazing angles  $\theta'_n$  become very small. This makes a very long cross point relative to an optical wavelength. The deflection electrode is 1 cm long whereas the wavelength,  $\lambda$ , is 0.00006328 cm. The ratio is about 15800 to 1. With a light wave only about 0.0001 cm long, a microcircuit sized cross point might have been expected. This is not so because the electrically induced difference in the index of refraction of the  $\text{LiTaO}_3$  crystal shown in figure 5-47 is so small. This in turn causes very small grazing angles.

#### 5.4.2 Switch Performance

A 3 x 3 switch was made by depositing metal electrodes on a  $\text{LiTaO}_3$  sheet crystal. The electrode pattern shown in figure 5-51 gives an idea of the switch functioning and size. Dimensions shown are in mils. The voltages applied set the critical grazing angle greater than 1.8 degrees and so the angle between horizontal and vertical channels was 3.6 degrees. On this scale, a 100 x 100



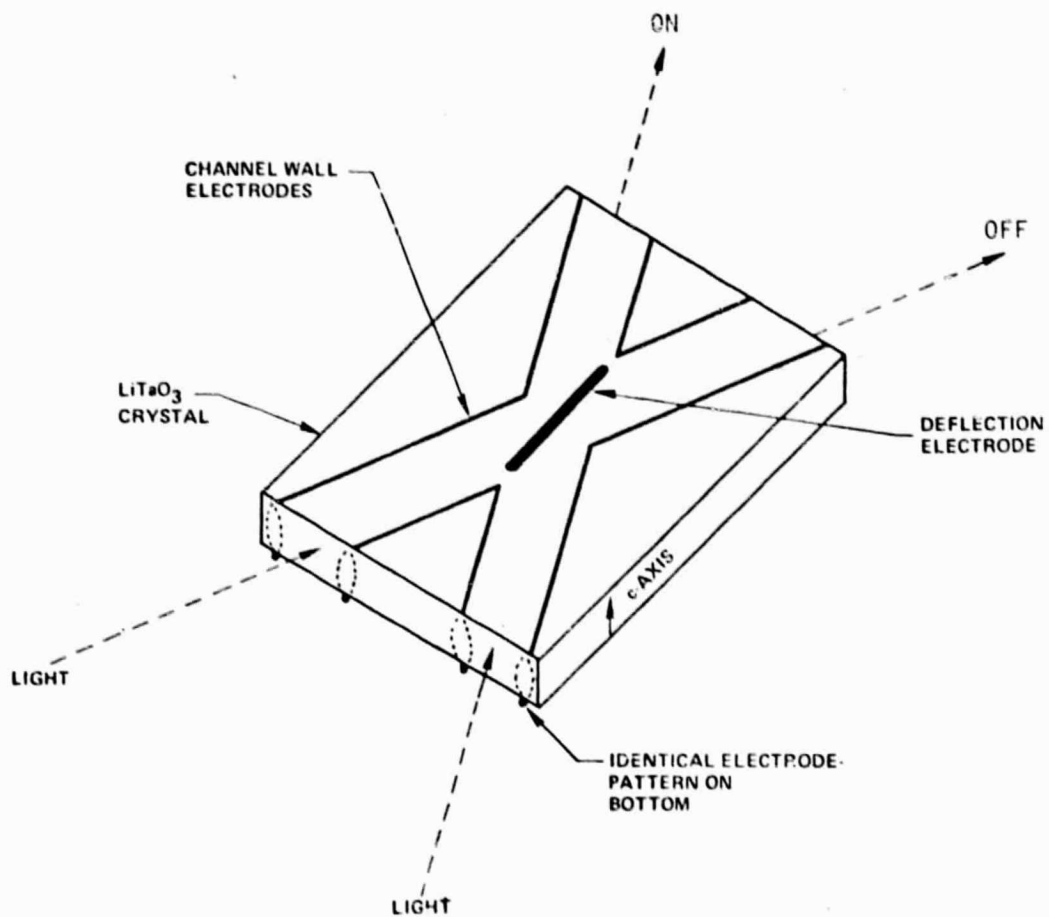


Figure 5-49. Optical Cross Point

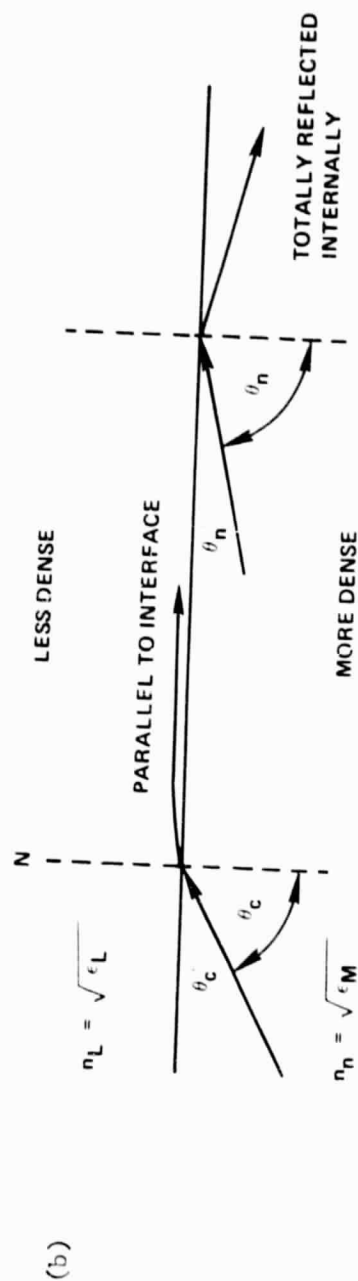
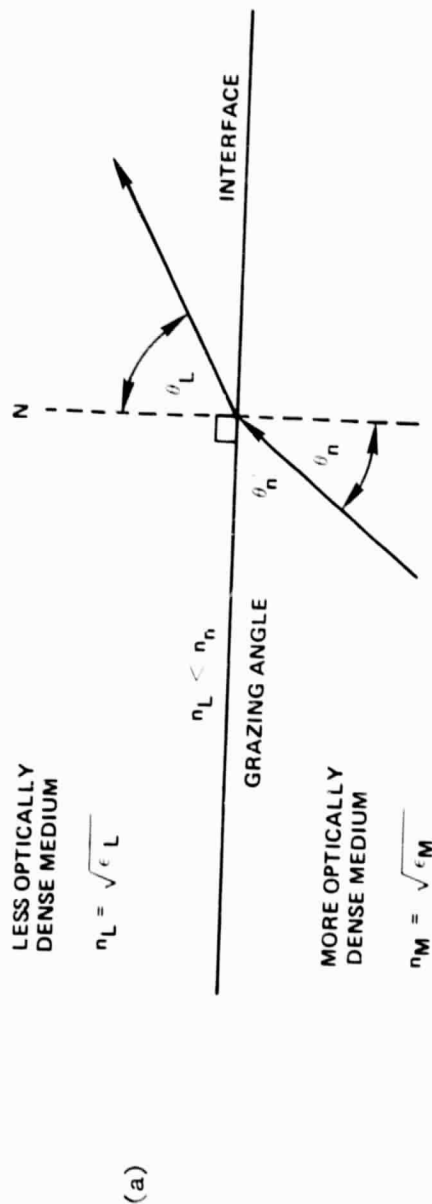
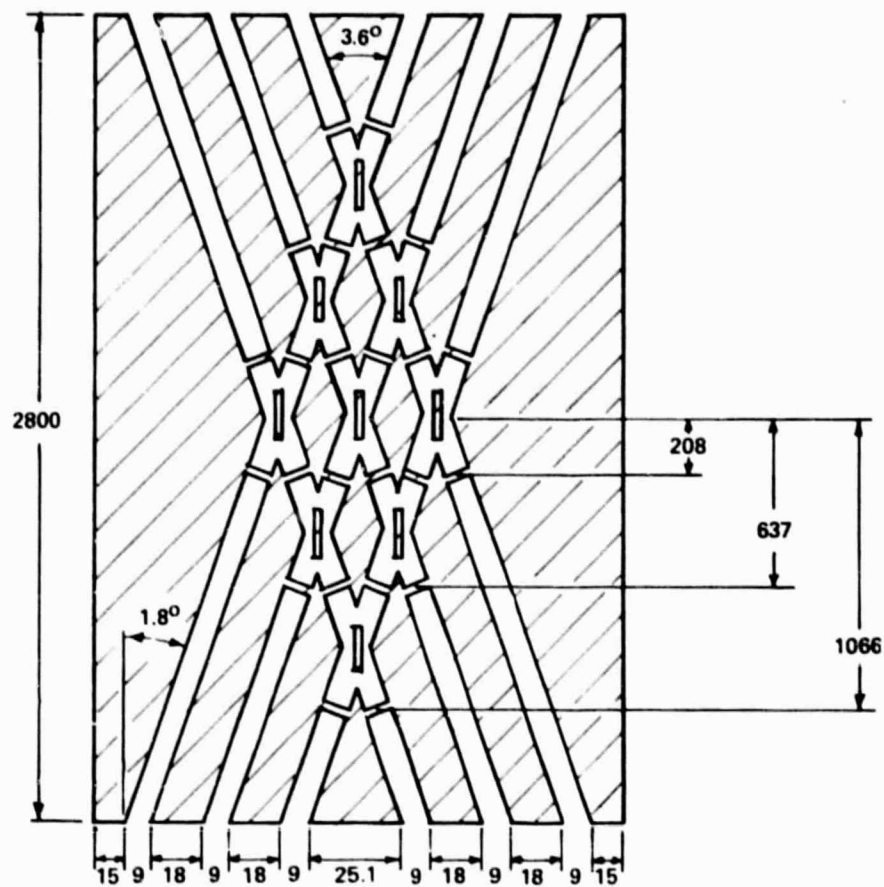


Figure 5-50. Wave Guiding and Switching by Internal Reflection



PHOTORESIST MASK 3 X 3 -1 FOR TOP-SURFACE ELECTRODE PATTERN OF CROSSBAR. SHADED AREAS REPRESENT METAL ON THE CRYSTAL. DIMENSIONS IN MILS.

Figure 5-51. Optical Switch Electrode Pattern

switch would occupy an area of 18,000 square inches if the insertion loss to isolation and signal to cross talk performance ratios are satisfactory.

The switch was driven with voltages between 400 and 500 volts. Little current was drawn due to the high resistivity. Switching times were short but not measured. They should be very fast (about a few nanoseconds) because the d.c. electric field response of the crystal is rapid and it does not take the field long to penetrate the 80  $\mu$  thick crystal. The limitation would therefore be the ability to switch 500 d.c. volts rapidly in a high impedance circuit.

The total measured insertion loss of the 3 x 3 switch is 19 dB and the isolation is 8 dB. An earlier 2 x 2 switch had an isolation of 16 dB. Improvement of the switch requires materials with more favorable characteristics. These would simultaneously reduce switch size while increasing isolation.

#### 5.4.3 Optical Switching Conclusions

Optical switching is not competitive at its present state of development. Its advantage lies in the ability to switch in an optical medium for transmission systems which are themselves optical without the necessity to transduce from optics to electronics and back again for switching. The reverse presently applies to a microwave satellite communications system. Unless there is an outstanding advantage, there is no point in leaving the microwave medium to switch at optics. Optics does raise the hope, since light ray cross sections can be so small, of considerably reducing the size of the switch. Optics also requires only a single layer construction for a true crossbar configuration. However the hope of size reduction is not presently realized. It may be that improved crystals will be found in the future. Therefore optics should be looked at again in five to ten years. The performance does not presently warrant the penalty of the extra transducing equipment.

#### 5.5 MESFET SWITCHING TECHNOLOGY

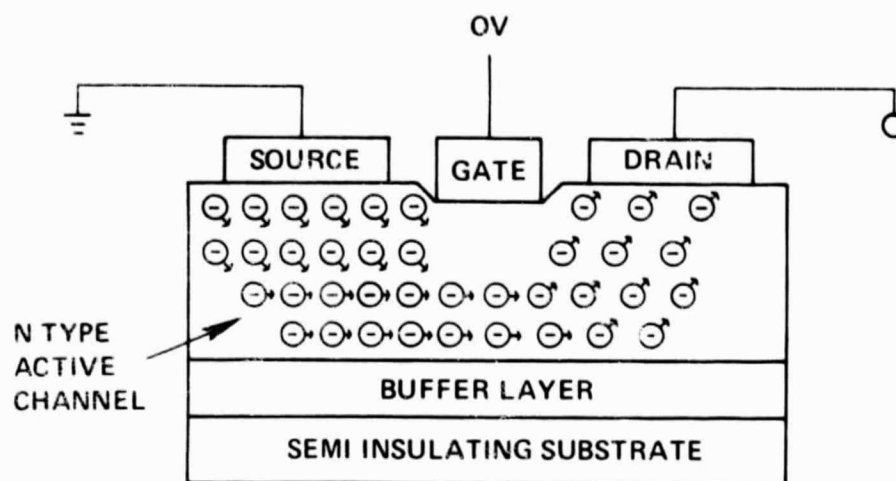
A most promising device for use in a large microwave switch is the MESFET. It is presently the highest frequency solid state active device in use for both low noise and high power microwave applications. Its noise figure is not as good as the cooled paramp, but its performance cost ratio is unexcelled. Noise figures of 3.5 dB and powers to 10.0 W at frequencies up to 9 GHz have recently been obtained with FETs. Since any device which amplifies and

oscillates can switch, attention is now turning to utilizing FETs for microwave switching. Part of the reason for the high speed of FETs is the avoidance of P-N junction barriers which tend to limit speed in traditional bipolar transistors. Only Schottky (metal semi-conductor) barriers are used in MESFETs, which are therefore very fast.

Another factor improving the speed performance of FETs has been the use of gallium arsenide (GaAs) which has three important advantages over silicon. First, its electrical resistance is about 1/7 that of silicon. Second, electrons can travel about five times as fast in gallium arsenide. Third, GaAs can be doped with chromium and iron to render it semi-insulating. This property allows the fabrication of almost totally non-conducting substrates, a property which silicon does not have. This substrate is needed for high frequency performance and facilitates fabrication of monolithic circuits [DiLorenzo and Schlosser, 1978]. Until recently, problems attendant on fabricating devices with GaAs limited development. GaAs is a compound and therefore its properties are more complex than silicon and it is more brittle and delicate to handle. Recent development of epitaxial growth technology capable of producing extremely thin layers, and advances in high resolution photolithography and lift-off processes to form very narrow gates, have made production of GaAs FETs practical.

#### 5.5.1 GaAs MESFET Configuration

Figure 5-52 shows a simplified cross section of a MESFET. There are two types of charge carriers used in semi-conductors, electrons, and holes. Electrons move much faster than holes in gallium arsenide so GaAs FETs use N type material. The FET electron current flows in an N type region which is highly doped with donors to have very good conductivity. It is injected into the region at the source electrode and flows to the drain electrode under the influence of a positive potential relative to the source applied at the drain. A control electrode called the gate can encourage or hinder the electron flow by virtue of its position between the source and drain if a small positive or negative voltage, respectively, is applied to it relative to the source. Because of its direct connection to the N type channel, the gate has a slight back bias when no voltage relative to the source is applied. The other two electrodes have a highly doped conductive pad between themselves and the channel to enhance electron injection and recovery. A buffer layer of intermediate conductivity is deposited between the N channel active layer and the insulating substrate to



1A-34,820

Figure 5-52. MESFET Arrangement

reduce thermal scattering of electrons in the channel by the substrate which would introduce noise and decrease channel conductivity.

The necessary layers are laid down by epitaxial growth, a process which allows the formation of thin films on the surface of crystals. These films have the same crystallographic properties as the substrate. The films must be thin, pure, uniformly deposited, and have very smooth surfaces. The level of doping must be changed rapidly without interrupting the growth of the various layers. The relatively thick chromium doped substrate insulates the device electrically, gives it mechanical strength, and provides a base on which the various layers are grown. The very lightly N doped buffer layer, which is 3 to 5  $\mu$  thick, isolates the active layer on top from the insulating layer below. Atop the buffer is the N channel active layer 0.2 to 0.5  $\mu$  thick, which forms the channel for electron flow. Finally, highly doped N+ pads are grown for good contact with the source and drain electrodes. Metal is alloyed to the N+ pads to form source and drain electrodes. The gate electrode is placed in a trough between source and drain directly in contact with the N type channel. Additional metal coatings are built on the electrodes to decrease resistance and improve reliability.

#### 5.5.2 MESFET Operation

The presence of a gate electrode between the source and the drain makes the MESFET a three terminal device capable of amplification and control. Electron flow is encouraged or inhibited by the electric field projected by the gate into the active conducting channel. Little current flows to the gate itself which exhibits, therefore, a relatively high input impedance. A series of curves of drain current versus drain voltage may be drawn with gate source voltage as a parameter. One such set of currents is shown in figure 5-53. The figure also shows a typical resistive load line with ON and OFF switching states. As can be seen from this discussion, there is a strong resemblance between an FET and a vacuum tube triode.

A fairly accurate theory exists for the operation of the MESFET. It is based on the depletion layer formed under the gate because of its direct metal to semi-conductor contact which is more abrupt than the source gate contacts. This depletion layer acts like an insulating region and constricts the cross section available for current flow in the active channel. The width of the depletion region depends on the voltage applied between the semi-conductor and the gate. In figure 5-54a, the gate is shorted to the source and a small drain voltage is applied. In this case, the depletion layer

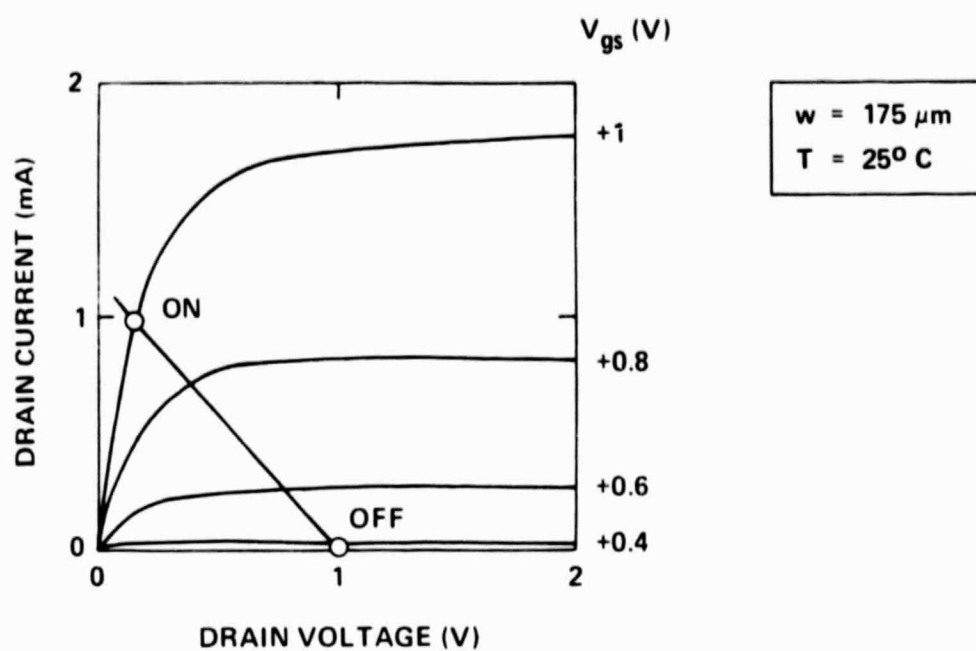


Figure 5-53. Drain Current vs. Drain Voltage



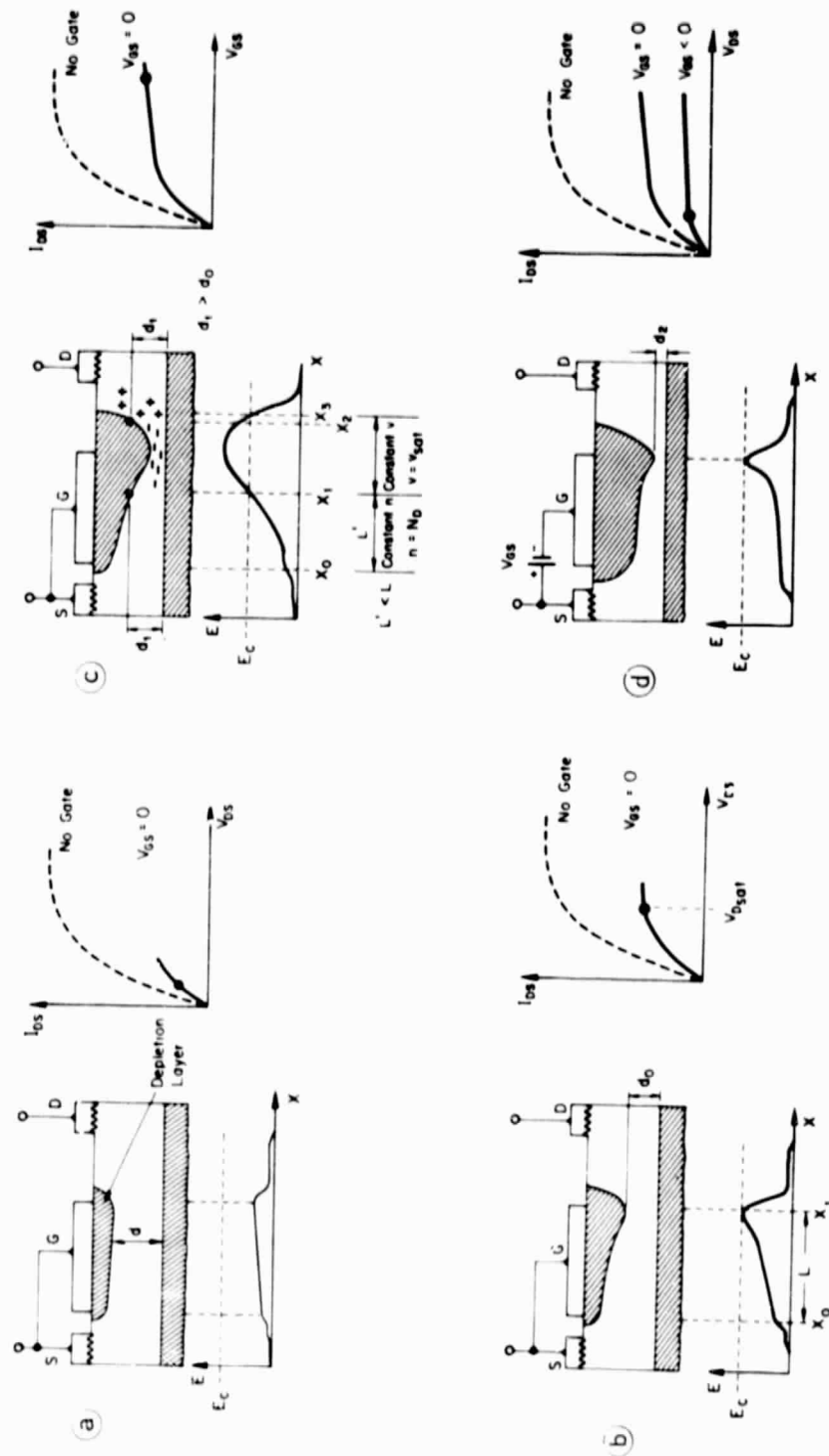


Figure 5-54. MESFET Operation

has a small width and the conduction channel is narrowed somewhat. The resistance between source and drain is increased over what it would be in the absence of the gate.

For low drain-source voltages, the conducting region behaves like a linear resistor as shown in figure 5-54b. Two effects then set in. The first, which is not dependent on the gate, occurs when the drain voltage is raised to a critical value,  $E_c$ , at which electrons reach a maximum velocity and current is maximized. The second, for any fixed value of drain voltage, is the increase of potential as the drain is approached in the channel. Because the gate itself is an equi-potential surface, the gate is increasingly back biased as the drain is approached. This results in a widening of the depletion region on the drain side of the gate, as shown in figure 5-54c and 5-54d, and causes the formation of a dipole layer as the drain voltage is increased beyond saturation. All of these processes result in a leveling off of the drain current versus drain voltage characteristic. The addition of a negative bias to the gate widens the depletion region, and causes the onset of the saturation effects earlier in the drain current versus drain voltage curves, thereby depressing them as shown in figure 5-54. These processes are described in detail in [Liechti, 1976] and they culminate in the family of curves of figure 5-53 which is a prototype design tool for the microwave engineer. The theory is accurate enough to produce a lumped element equivalent circuit for the MESFET shown in figure 5-55, which indicates the relationship of its physical elements to the circuit elements. Typical values of circuit elements are given in table 5-5. Circuit elements would be different for an FET optimized for switching.

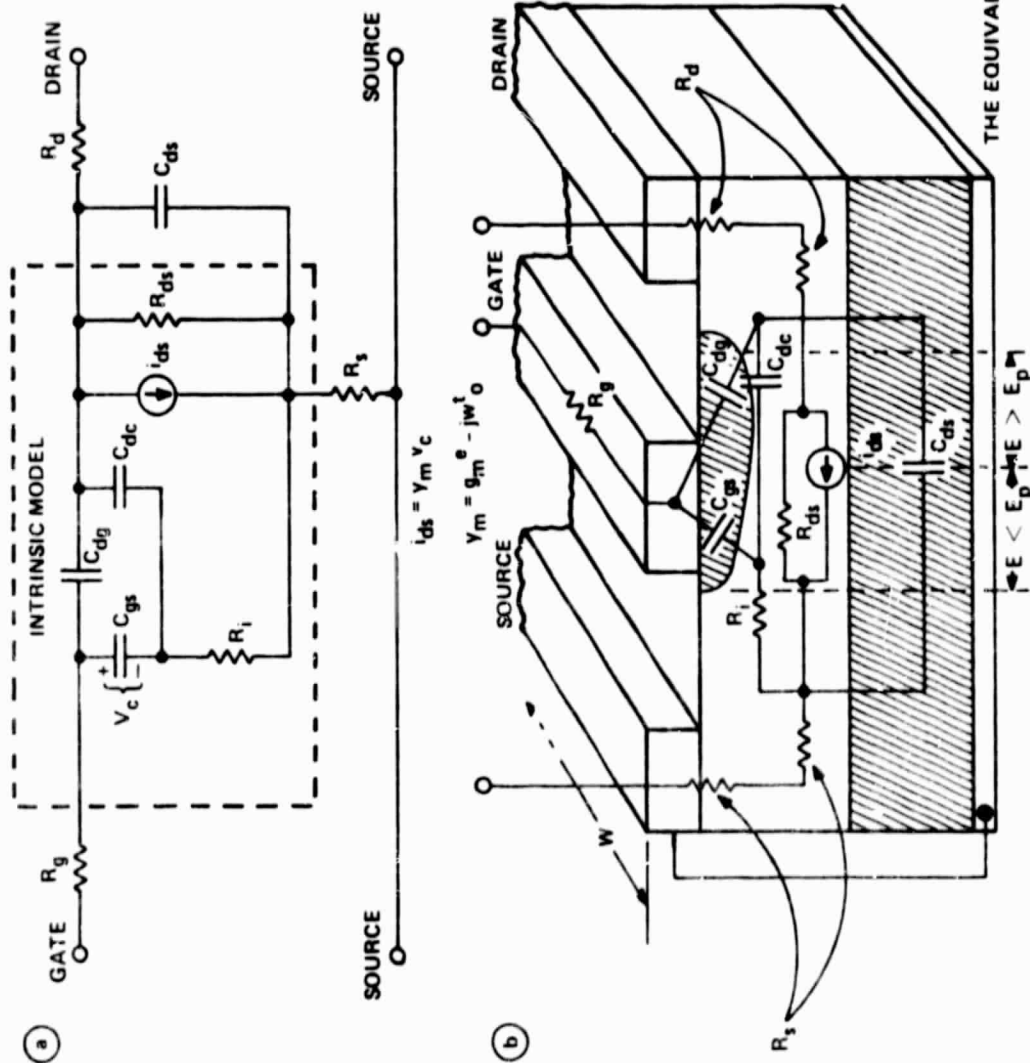


Figure 5-55. MESFET Principles

Table 5-5

Equivalent Circuit Parameters of a Low-Noise GaAs MESFET  
 with a  $1\text{ }\mu\text{m} \times 500\text{ }\mu\text{m}$  Gate  
 (HP Experimental,  $N_d = 1 \times 10^{17}\text{ cm}^{-3}$ )

$g_m = 53\text{ mmho}$	$C_{ds} = 0.12\text{ pF}$
$t_o = 5.0\text{ ps}$	$R_g = 2.9\Omega$
$C_{gs} = 0.62\text{ pF}$	$R_d = 3\Omega$
$C_{dg} = 0.014\text{ pF}$	$R_s = 2.0\Omega$
$C_{dc} = 0.02\text{ pF}$	$L_g = 0.05\text{ nH}^*$
$R_i = 2.6\Omega$	$L_d = 0.05\text{ nH}^*$
$R_{ds} = 400\Omega$	$L_s = 0.04\text{ nH}^*$

d.c. bias

$V_{ds} = 5\text{ V}$   
 $V_{gs} = 0\text{ V}$   
 $I_{ds} = 70\text{ mA (saturated)}$

---

\*Contacting inductances of the test fixture in series with  $R_g$ ,  $R_d$ , and  $R_s$ , respectively.

### 5.5.3 MESFET Switching Applications

Two microwave switching applications of MESFET appear feasible. In the first, [Gaspari and Yee, 1978] studied switches in the configurations shown in figure 5-56. They found the parallel form difficult to use because resistance of the FET used was 30 to 40 ohms, which made turnoff isolation poor. The gain and series gain and series-series configurations provided good isolation but were not chosen in order to minimize bias power based on the then current state of FET development.

The series configuration was chosen. It makes use of saturation at a gate voltage of 0.0 volts, and pinchoff at a gate voltage of 4 volts, and provided an inherently broadband untuned response with effectively zero d.c. bias power, since the gate source Schottky diode was always reverse biased and no external voltage was applied to the drain. Measurements were made with an older FET, a Plessey GAT-2  $2\mu$  gate device. The swept broadband untuned and tuned insertion losses and isolations achieved are shown in figure 5-57. When Gaspari and Yee used the circuit model of the switch, they decided that they could inductively resonate the source to gate and gate to drain capacitances to increase the isolation. Figure 5-57 shows the isolation with and without the shunt inductor. Resonance improves the average isolation from 10 dB to 21 dB. Tuning for resonance was accomplished by the addition of a fine wire in parallel. Gaspari and Yee state that the more widespread  $1\mu$  gate geometry is capable of providing even better isolation in a single stage. The authors built an 8 x 8 switch using the single pole eight throw fan-out/fan-in architecture. This switch provided the isolation shown in figure 5-58. Switching speeds of 1 ns were observed by Gaspari and Yee.

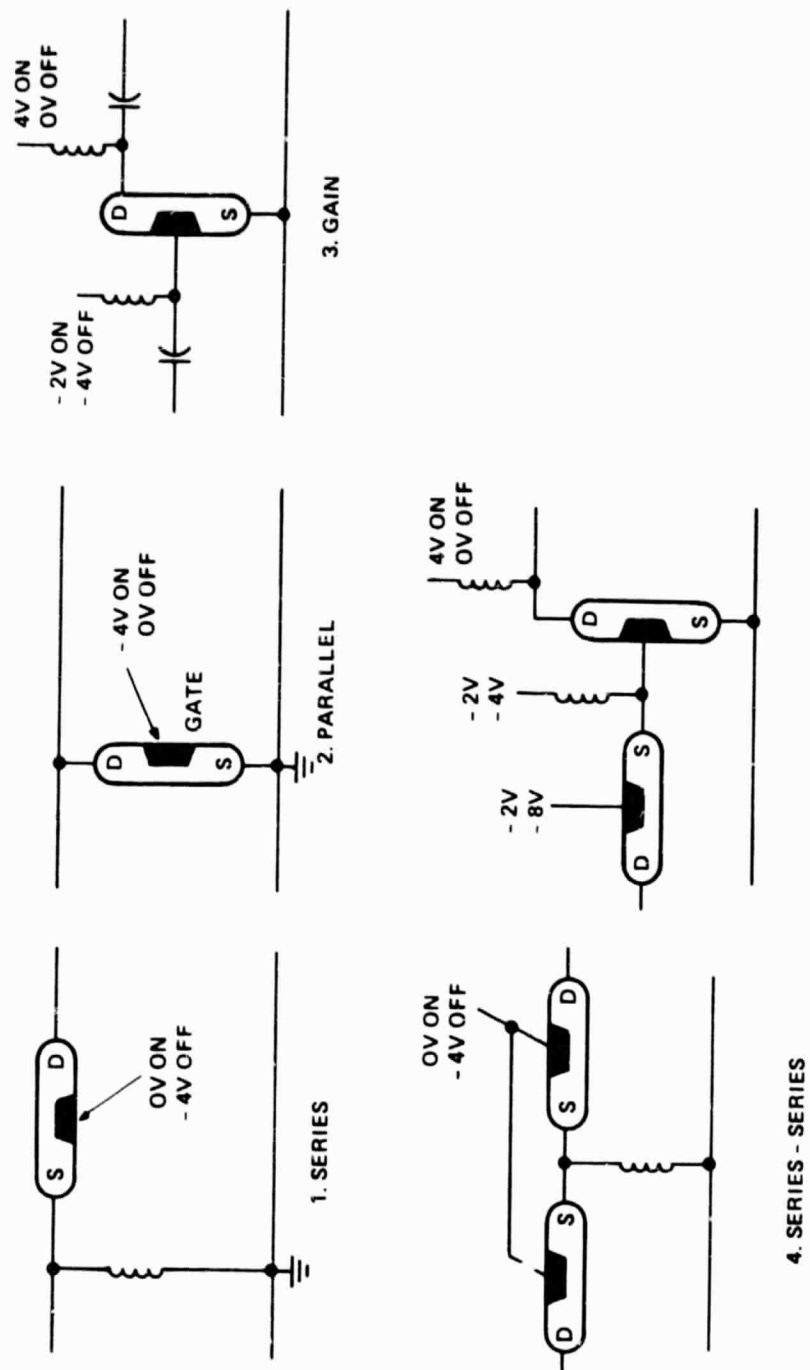


Figure 5-56. FET SPST Configurations

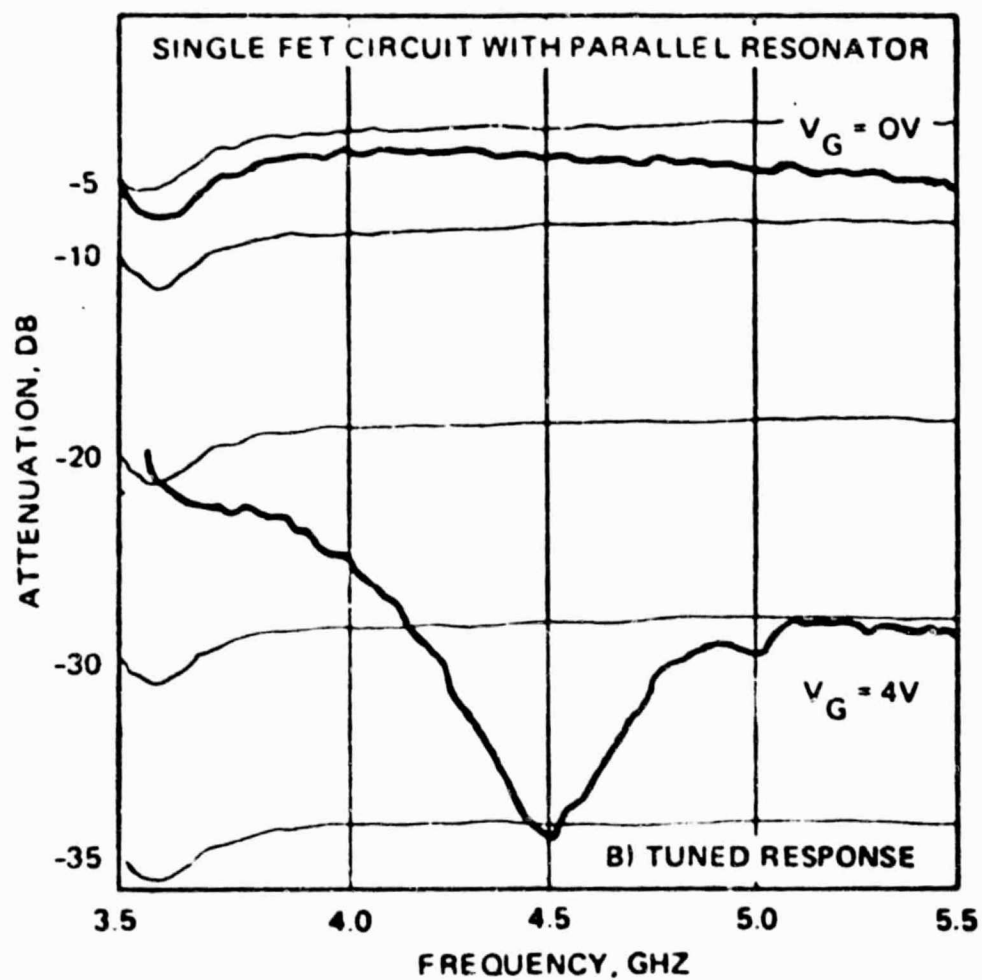
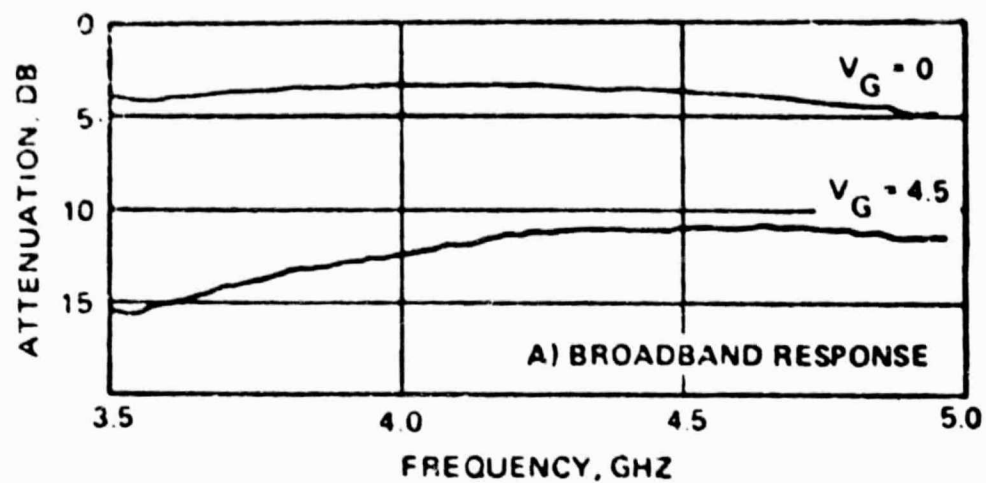


Figure 5-57. Performance of Plessey MESFET as a Series Switch

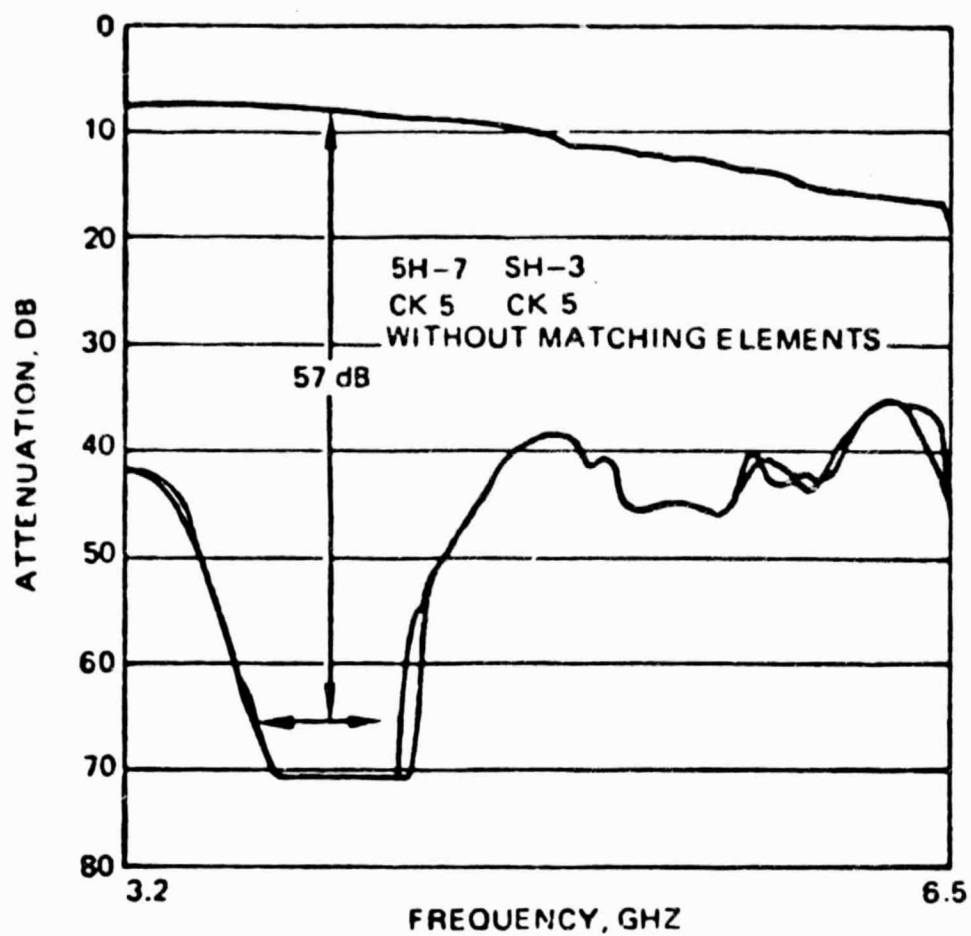


Figure 5-58. Performance of 8 x 8 MESFET Switch



Gaspari and Yee were reluctant to use the series gain configuration because a single gate FET would then require bias power. They exploited the high input impedance of the gate source port. [Goel and Wolkstein, 1978] employed a dual gate MESFET to circumvent this problem. A view of a dual gate schematic is shown in figure 5-59. One gate, G1, was used for microwave amplification and the other gate G2 for independent video control.

The amplifier exhibited a maximum gain of 20 dB, a gain flatness of 0.75 dB, a noise figure of less than 7 dB, group delay of 0.3 ns over a 4 to 8 GHz band, and provided an output power of 13 dBm at its 1 dB compression point. One of the most striking characteristics of the amplifier was its enormous dynamic range: by varying the gate 2 voltage, the gain/loss performance of the unit over the entire frequency octave can be changed from a loss of 40 dB to the maximum gain of 20 dB while maintaining the gain flatness over the band. This range was achieved for a swing in gate 2 voltage from +1.0 V ( for the gain of +20 dB) to -2 V (for the 40 dB insertion loss). This performance is shown in figure 5-60.

With RF input applied to the first gate, the second gate can be used for high speed switching. An experiment was performed to measure the amplifier response to a square wave in which an input pulse having a rise and fall time of 200 ps and a pulse duration of 1.5 ns was applied to gate 2. This pulse drove the second gate of the amplifier from 0.0 to -0.7 V, thus changing the gain of the amplifier by about 18 dB, while causing less than 100 ps increase in the rise and fall times.

#### 5.5.4 MESFET Switching Conclusions

FET switching is new. The MESFET, which is inherently simple to make, is not yet optimized for the switching function. Suggestions for improving its performance as a control element have been made by R. V. Garver, the pioneer in solid state switching [Garver, 1979]. He discusses measures such as doping profiles to reduce losses.

FET has a number of advantages for switching:

1. The FET biasing input is inherently isolated from the RF line and requires minimum block choke separation.
2. There is a possibility of achieving insertion gain rather than loss with the FET.

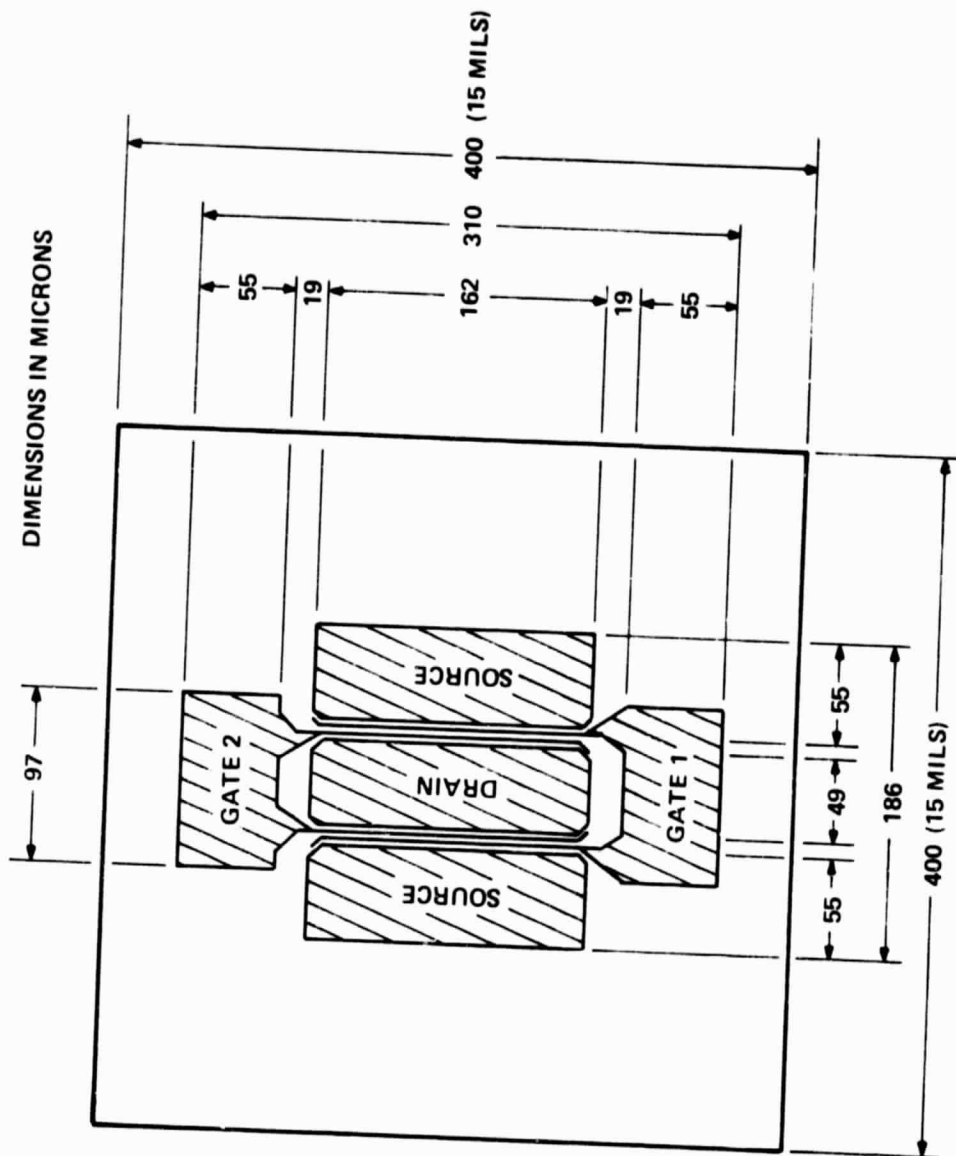


Figure 5-59. Dual Gate FET

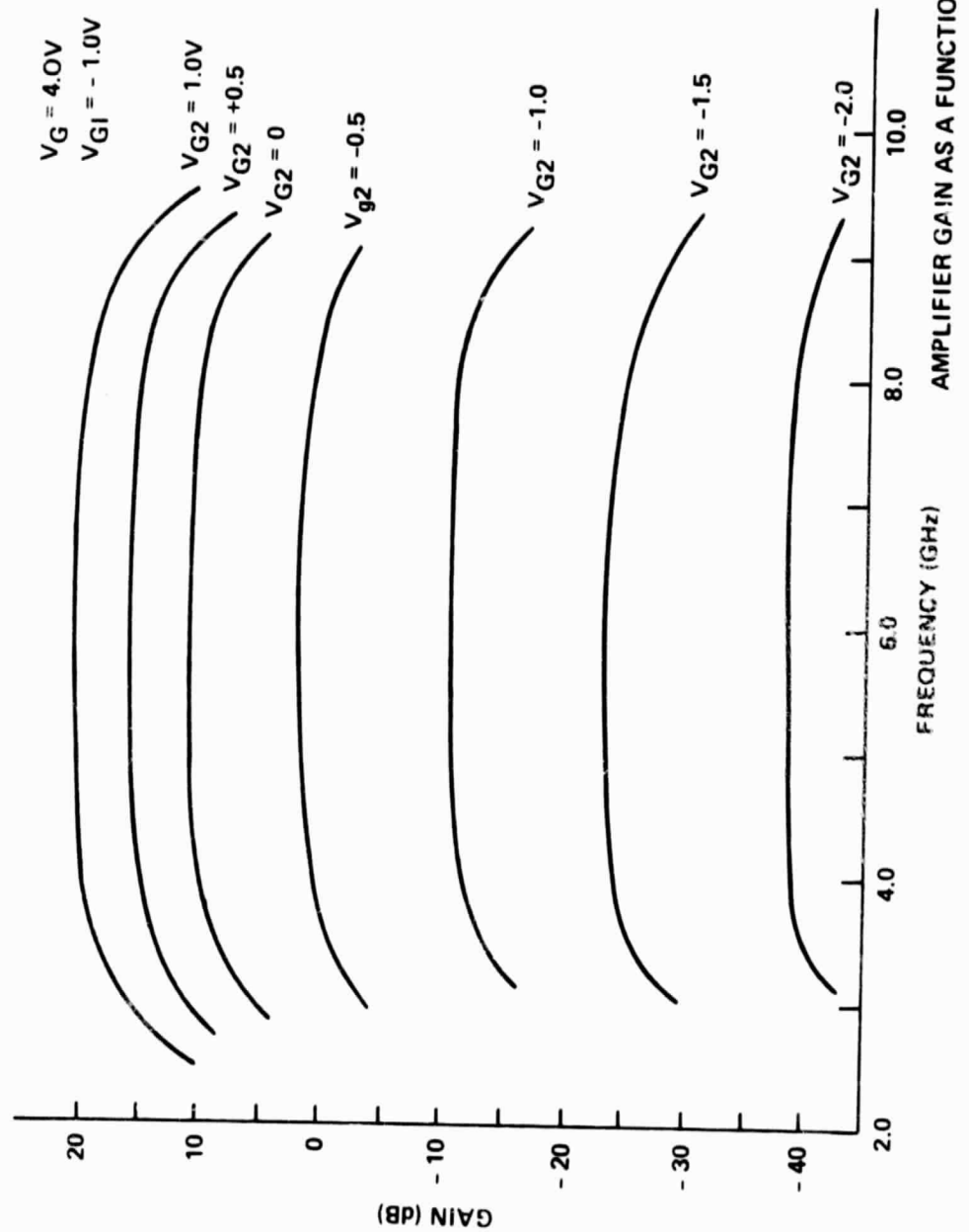


Figure 5-60. MESFET Amplifier Performance

3. There is a possibility of achieving essentially very low d.c. power dissipation in driving the switch to its ON and OFF states.
4. The FET provides a distinct speed advantage over the PIN diode.
5. The FET can be driven by the low voltage logic levels, thus eliminating the need for driver circuitry.
6. When used as an amplifier for switching, the FET does not have two modes of failure (i.e., open or short); it will fail only one way (i.e., as an amplifier).
7. Switching might be done at a higher frequency than 3 GHz.

The FET is so promising that research effort should be expended to optimize it for switching.

#### 5.6 SWITCHING SUBSYSTEM CONSIDERATIONS

Broadcast capability, memory, control, and reliability are all considerations in designing a switching subsystem.

##### 5.6.1 Broadcast Capability

It is desirable to provide a broadcast capability in the switch, that is, to be able to transmit a signal from any input to a number of outputs simultaneously. It is also desirable that this should include any or all outputs in the switch. Certain architectures already provide this capability. These include the fan-out/fan-in with couplers, and the cross bar architecture, again with couplers. The fan-out/fan-in architecture with single pole multiple throw switches has connectivity for broadcast but encounters an adverse input impedance when all series diodes of a SPMT microwave switch are simultaneously forward biased. This is a valid reason for favoring the first two architectures. Completeness requires mention of the wraparound concept for broadcast. This concept has more important uses in redundancy for reliability, but serves to introduce the concept.

The essence of the wraparound concept is the provision of one or more extra output lines which are accessible from all input lines. An individual extra output line is then routed or wrapped around to perform an additional function such as broadcast. This is

shown schematically in figure 5-61. The last output line is an extra available to all input lines. This line is brought out to an amplifier which boosts the signal so that it can be divided to serve many outputs. The output line is then wrapped around the switch so as to cross all the output lines which are interrupted by individual SPDT switches. These switches connect either the regular outputs or the broadcast output to the lines of the downlink transponders indicated in figure 5-61. Circuit elements not shown provide matching for the amplifier and the output lines.

#### 5.6.2 Memory Control

Initial consideration has been given to memory requirements on the satellite for switch control. The objective is to minimize the computer complexity aboard the satellite. However, a compromise is necessary for switch control because if all switch instructions were sent from the ground too much bandwidth would be required.

Assume that a  $125\ \mu\text{s}$  frame is repeated over and over again. There are approximately  $125/2$  or 63 different switch states in one frame time each of which represents an estimated change in information of 560 bits. Therefore, the total informational change needed for a frame is  $560 \times 63 = 35,280$  bits. If this information were sent continuously from the ground every  $125\ \mu\text{s}$ , the channel would have to transmit 280 Mbps. If we take advantage of the fact that a frame is repeated over and over again, we can store these bits on board the satellite and update them at a slow rate from the ground. A 35,280 bit store on board the satellite is an excellent trade in return for a 280 megabit channel down to earth from a satellite. The store is a memory of less than 4.5 K bytes.

It is expected that the connectivity sequence of a given frame will persist for a relatively long time. Therefore, the master station on the ground can send update information concerning switch states at an estimated 200 bps rate. This is received, buffered, aboard the spacecraft, and written into the cyclical non-destructive readout memory. The switch control channel receiver, and its buffer storage, comprise a subsystem known as the switch command demod subsystem. Its output is provided to the switch control system shown in figure 5-62.

The information for changes in a  $2\ \mu\text{s}$  slot arrives on the command demod line at the lower left in figure 5-62 as a succession of  $(K + B)$  bit words. The K field contains the address in the switch control memory for the B bits being received. The  $(K + B)$  bit word is received in a buffer register and, at a strobe command from the timing control, is transferred to the write register. Here

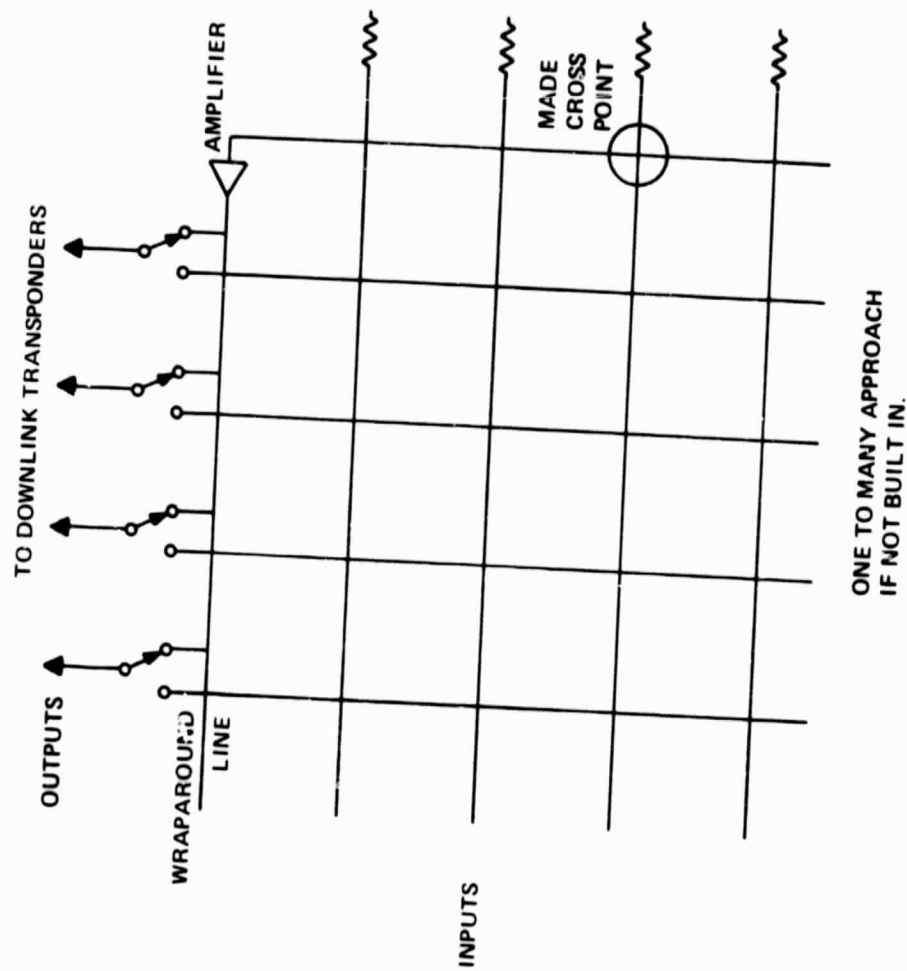


Figure 5-61. Wraparound for Broadcast

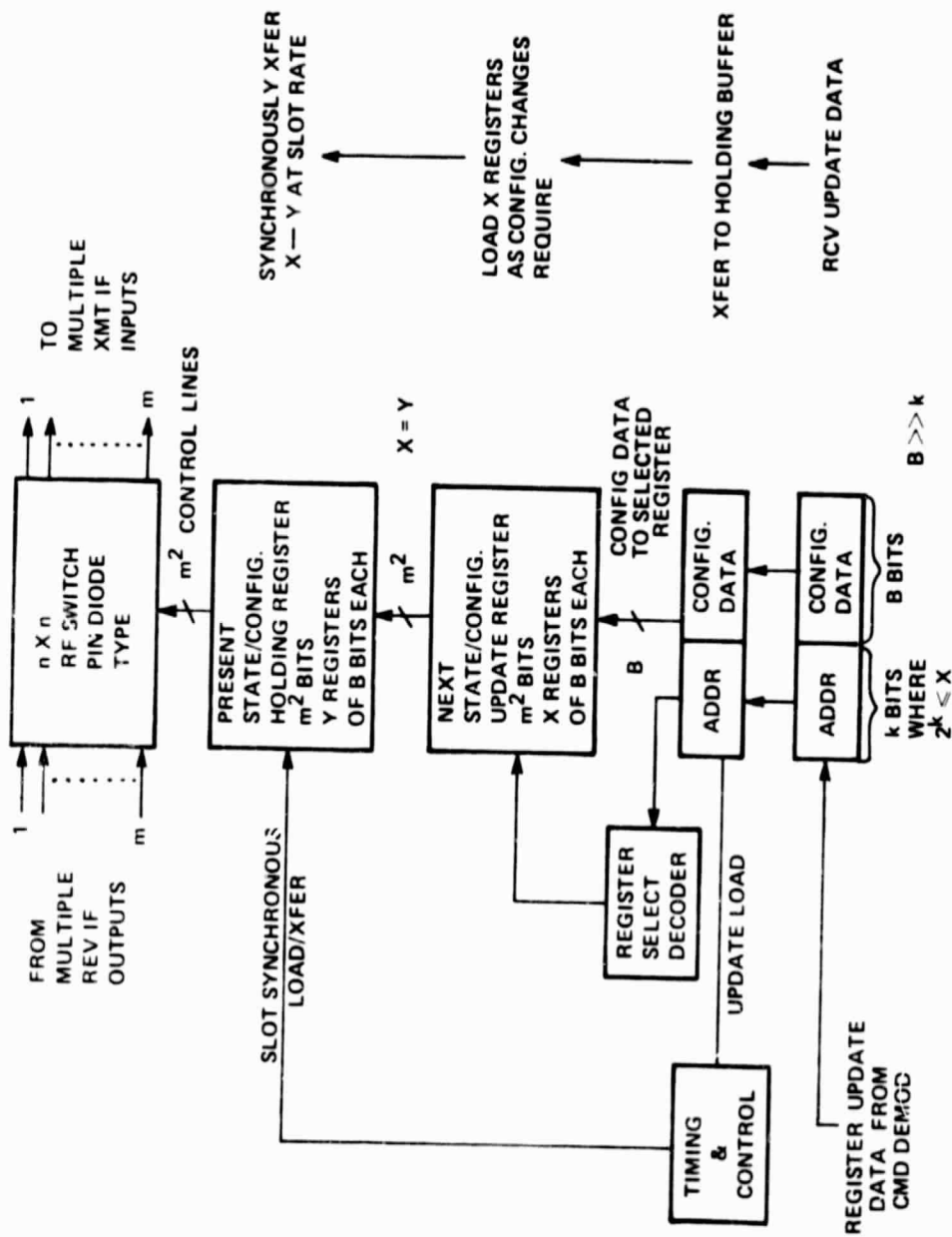


Figure 5-62. RF Switch Control Memory

the K bit field is decoded and is written into a buffer switch memory large enough to command every cross point in the switch at the command of another timing strobe. Within the  $2\ \mu\text{s}$  slot, all the changes needed for the next slot are written into buffer memory, and the buffer memory is completely dumped as a 10,000 bit parallel transfer to a switch control memory which is identical to the buffer in all respects. All the bits occupy the same positions in the switch control memory that they did in the buffer. Shortly after the dump at the start of the new slot, the switch control memory energizes 10,000 lines which go to the switch cross points in accordance with the one/zero patterns in the control memory. Each of the 10,000 lines has a devoted logic circuit which compares a new bit status with the previous status. If they are the same, the voltage on the line is not allowed to change. No disturbance is experienced by channels intended to operate through the new slot. Those lines which change logic state proceed to the drivers operating the devices at the switch cross points. The drivers receive the logic voltages and draw on the power supply for the energy to form the spikes and pulses needed to operate the solid state devices at the cross points. If the solid state devices can be operated by the logic levels directly, then the 10,000 drivers and the energy they use become unnecessary.

A straightforward logic control system for on-board operation of the switch has been described. The units do not have an unusual number of components or weights. The system requires one or at most two 4.5 K byte memories and two 10 K bit memories, plus some buffer registers. There are some serial data transfers at a high data rate of 280 Mbps and register transfers and memory dumps which would need to operate about 20 times slower. This performance should be achievable, and no R&D effort is deemed necessary to obtain it. Further design might suggest an even simpler system.

### 5.6.3 Reliability

There are 10,000 solid state cross points in the TDMA 100 x 100 switch. During a seven-year postulated satellite life time, some of these will fail. The failure of a cross point need not mean the failure of the switch to perform its function. Failure of the switch will be defined as the inability of the switch to make a connection between two cities. Assume that the failure of a cross point causes only the loss of service of that one cross point. That is, if nothing else is done, the interconnection between city L and city Q has been lost, and only that interconnection. Both L and Q can still interconnect to the other 98 cities. A failure of the switch will be defined as the inability to make a connection between L and Q, and the reliability goal of the switch will be that there



should be a 99.99 percent probability of no such failures in seven years, i.e., only one interconnection failure in that time.

In order to meet such a goal, switch architecture and circuit design should be chosen with a view to eliminating critical point catastrophic failures. That is, the failure of a solid state element should lead to no worse than the loss of a single cross point, and not to the outage of several cross points. An example of a critical failure is the loss of a device in a fan-out using SPMT switches, as in figure 5-8. In this case, an entire second rank SPMT switch connecting to as many as 10 cities would become useless. Hence fan-out and fan-in architecture based on SPMT switches is not recommended from a reliability point of view.

Assume that the switch has been designed so that catastrophic failures cannot occur. If several solid state devices are going to fail, the goal of 99.99 percent probability of no switch failure in seven years can be achieved nevertheless by redundancy through wraparound. This concept has two parts. The first is the illustration in figure 5-63, which shows the wraparound of an extra output line so that it connects to the input of an extra input line. The redundant wraparound works as follows: Suppose the cross point connecting city 2 to city 3 fails and city 2 has traffic for city 3. A routine test will show to the control station that cross point 2, 3 has failed. This is duly entered into the control computer on the ground which sends up a connection change in the usual way on the 200 bit uplink switch command channel. The change would route the signal to the cross point between input line 2 and the output wraparound line, i.e., cross point 2, W1. The failure of 2, 3 cannot interfere with the operation of any other cross point, so the signal can reach 2, W1. The signal then follows the wraparound line to the extra input line to the switch, input line 101. It then proceeds across the input line until it reaches output line 3 at cross point 101, 3. This cross point is closed and the signal exits on line 3 to its destination. Thus, a connection is made between city 2 and city 3 despite the fact that cross point 2, 3 has failed. This ends part one of the concept.

It should be pointed out that the extra output line and extra input line connected by wrapping around the switch are capable of serving any failed cross point in the switch. This back-up is attained at the cost of 200 extra cross points. What if two cross points fail? If the two failed cross points are never closed at the same time, then the single extra output and input wraparound lines will suffice by definition. If they do close at the same time, then two extra output lines and two extra input lines are needed. It is desirable to reduce the number of extra input and output lines we

1A-34,016

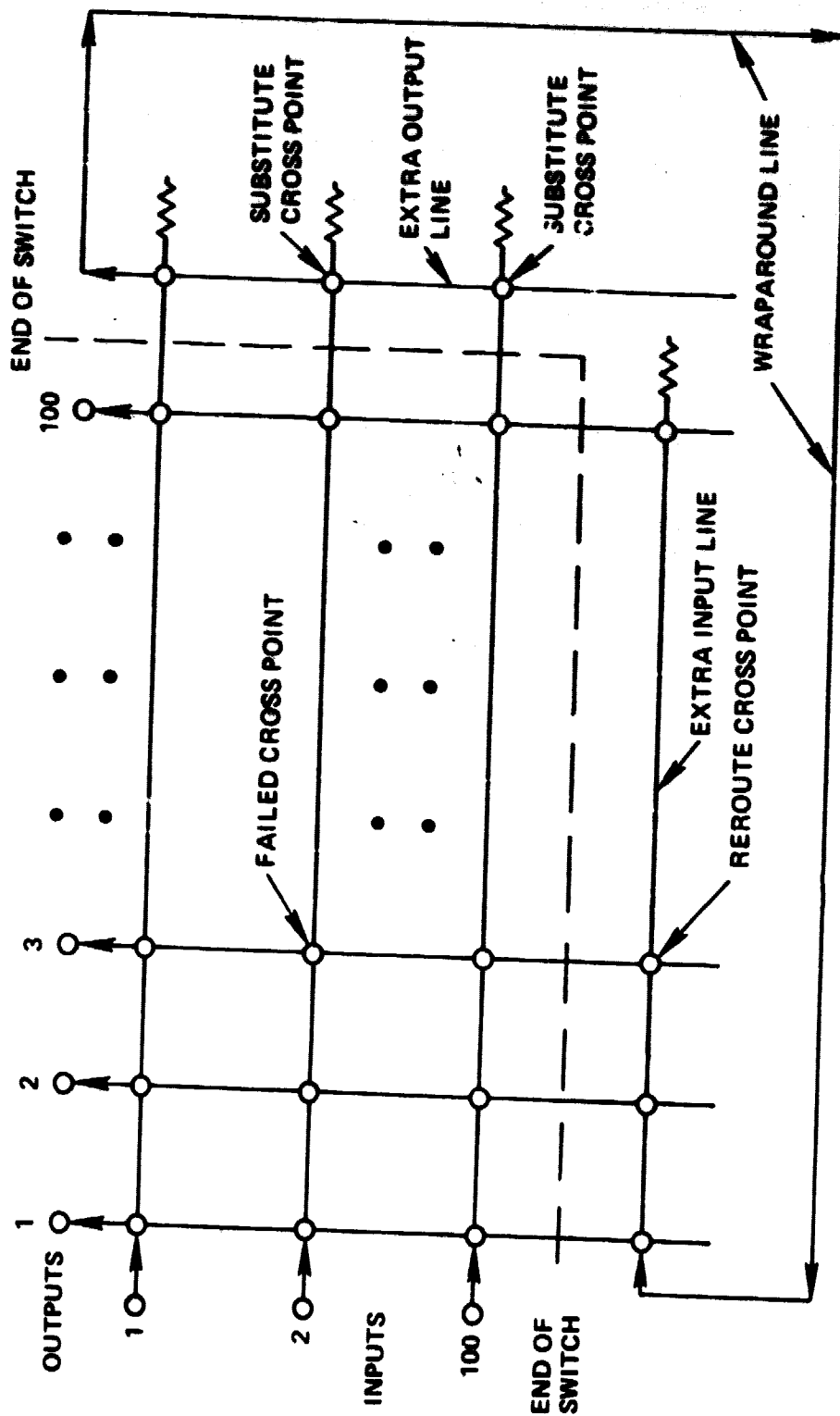


Figure 5-63. Wraparound Redundancy

C-4

add to the switch, since the number of cross points increases in multiples of 200. This begins part two of the concept.

The number of extra input/output line pairs plus wraparound lines needed is given by the number of failed cross points required to be connected simultaneously. This is given by  $10,000 \times$  the probability of failure of a cross point in seven years  $\times$  the probability that a failed cross point will be on at the same time as another failed cross point. An estimate of this number is made for an FET coupler type switch based on the following assumptions:

1. The failed cross points are evenly distributed throughout the switch.
2. A diagnostic testroutine is periodically run by the ground control station to determine the health status of the cross points.
3. At any time, 100 cross points are connected.
4. MIL-Handbook-217-B failure rates apply, and FETs are considered.

MIL-Handbook-217B was consulted and used to calculate the number of FET failures in seven years using the following values for constants in the failure prediction calculation for  $L_p$

$$L_p = L_b (\pi_E \times \pi_A \times \pi_Q \times \pi_{S2} \times \pi_C)$$

where the values used were

$$L_b = 0.017 \text{ or } 0.088 \text{ (30 or 100 percent voltage)}$$

$$\pi_E = 1.0$$

$$\pi_Q = 0.2$$

$$\pi_A = 5.0$$

$$\pi_C = 1.0$$

$$\pi_{S2} = 0.3$$

The values computed for failures per  $10^6$  hours were 0.0051 for 30 percent of rated voltage and 0.0264 for 100 percent of rated voltage. Then 10,000 FETs would experience three failures or 16 failures, respectively, for 30 or 100 percent rated voltage operation in seven years (61,320 hours). The square root of the

product is found to be seven failures. For ease of calculation, this number was increased to 10 failures expected in seven years.

Next, the assumption was made that the failures are equally distributed throughout the switch. That is, the failure density is 0.001 everywhere. For normal operation, 100 switches are closed at any one time. Among those 100 switches there is  $0.001 \times 100 = 0.1$  failed switch. We may state this as: We expect to find one failed switch in 10 randomly selected groups of 100 switches. This seems to imply that we can always guarantee that the message can go through if we provide one wraparound input/output pair with 200 extra cross points. This is statistically 10 times the protection needed. Events are not linear however. While not probable, there is always a chance that a group of 100 cross points might have two failures. The probability depends on the distribution about the mean for the locations of the failed cross points. This is unknown.

What can be said is that there is an a priori probability of 0.001 that any of the 10,000 cross points in the switch will fail. Extend this to the 200 cross points in the wraparound; the probability that the wraparound itself might fail in just the way that would prevent the signal from reaching its intended destination is the joint probability of the two independent events of failures of particular cross points in the switch and wraparound. The joint probability for such independent events is the product  $0.001 \times 0.001$  or 0.000001. There are 200 ways this event could happen. The probability is 0.0002 that it might. One wraparound provides a probability of success of 0.9998. Consequently it is wise to add a second wraparound. This will both increase the probability of success by changing it to 0.99996 and help to cover the unlikely case where two failed cross points in the switch are included in a group of 100.

One more precaution should be taken to safeguard the switch. The ground control station follows which cross points have failed by telemetry. It can always permute the slots in a frame in such a way as to prevent the inclusion of more than two failures in a group of 100 operating cross points.

## **5.7 SWITCH STUDY CONCLUSIONS**

As a result of the study discussed in this section, certain specific conclusions have been reached. These are:

1. A future TDMA multibeam satellite communications system can offer connectivity and enormous communications capacity to the United States. For metropolitan area populations, this system will need to be supported by a large switch in excess of 75 x 75 cross points. In order to identify problems, 100 x 100 is an appropriate model.
2. Tentative specifications are listed in section 5.1.2.
3. Coupler cross bar architecture illustrated in figures 5-17 and 5-19 is superior from the point of view of minimum weight and volume, broadcast capability, and reliability.
4. Phased arrays are too expensive and have too many circuit problems to be considered for switching applications. They could provide another practical service, temporarily positioning a beam in some area (for example a disaster area) which does not usually generate high communications traffic. This is not switching. The signals received or transmitted would be routed by the switch.
5. Ferrite devices are too slow, too heavy, and require too much power to be useful in a large switch.
6. Optical switches are not presently smaller than microwave switches and may not ever be so because of limitations on the ability to change the index of refraction of optical materials. They are fast, but require biasing of up to several hundred volts for operation. Signals would have to be transduced from microwave to optics and back to microwaves in 100 channels, which increases component expense.
7. PIN diodes are a mature technique, and can do the job. There have been improvements in drive power consumption but drivers will be required. PINs are no longer the fastest elements and they require more circuitry both microwave and biasing than do FETs.

8. FETs can potentially exceed PINs in:

- Switching Speed
- Required switching power
- Elimination of drivers (only logic needed)
- Suitability for monolithic arrays
- Ability to amplify

The last capability reduces losses and provides a more definite failure mode than PINs which can fail either open or short. FETs are so promising that they will be the best devices if optimized.

9. One of the most difficult areas connected with the switch is how to lay out its design so as to provide necessary microwave signal, control, mounting, and thermal cooling paths, and so that the switch can be assembled and maintained before (perhaps even after) launch. A common question asked was how will the switch be assembled? This is because normal three dimensional wiring techniques used in switching at lower frequencies for signals are no longer applicable at microwaves.

## 5.8 RECOMMENDATIONS

### ● Design Study

In view of the problem of switch layout at microwaves, a switch design study should be carried out. It should determine layout, packaging, and construction practice for a microwave cross bar (function) switch of unprecedented (100 x 100) size. The study should cover routing of microwave and control signals, methods of cooling, interconnection, fabrication, and assembly. The study should assume use of devices, architecture, and transmission media recommended in this report. The study should cover logic control, reliability and one-to-many broadcasting.

- Media

The switch should use coupler cross bar architecture with feed throughs realized in microstrip. This configuration affords the smallest size and weight and provides the highest reliability.

- Devices

There are two usable devices, PINs and FETs. FETs are so promising that, despite needed optimization for switching, they are the preferred device. A development program should be instituted with recognized FET suppliers to develop optimum FETs for switching.

Optical switching should be reviewed in about five years to see if a breakthrough makes it competitive considering the need to transduce into and out of the optical medium.

Phased arrays and ferrites are unsuitable.

- Switch Size

A 100 x 100 microwave cross bar switch is a necessary and useful element in a TDMA satellite communications system. The size, 100 x 100, is justified in view of the 73 municipalities in the U.S. with more than a half million population plus the possibility of using of several beams for the largest municipality.

## REFERENCES

- Acampora, A. S., and Davis, B. R., "Efficient Utilization of Satellite Transponders via Time-Division Multibeam Scanning," B.S.T.S. Vol. 57, No. 8, October 1978.
- Armistead, M. A., Spencer, E. G., and Hatcher, R. D., "Microwave Semiconductor Switch," Proc IRE Vol. 44 pg. 1975, December 1956.
- Ayter, S., and Ayali, Y., "The Frequency Behavior of Stripline Circulator Junctions," IEEE Trans MTT, March 1978, page 197.
- Benes, V. E., "Mathematical Theory of Connecting Networks and Telephone Traffic," Academic Press, NY 1965, pp. 136-144.
- Betts, F., Temme, D. H., and Weiss, J. A., "A Switching Circulator: S. Band; Stripline; Remnant; 15 kilowatts; 10 microseconds, Temperature Stable," IEEE Trans MTT, December 1966, page 665.
- Bosch, B. G., "Gigabit Electronics - A Review," Proc IEEE, Vol. 67 No. 3, March 1979.
- Bosma, H., "On Stripline Circulation at UHF." IEEE Trans MTT, January 1964, page 61.
- Clos, C., "A Study of Non-Blocking Switching Networks," Bell System Tech J. Vol. 32, 1953, pp. 406-424.
- DiLorenzo, J. V., and Schlosser, W. O., "GaAs + FET = Improved Microwave Systems," Bell Laboratories Record, September 1978, p. 209.
- Fay, C. E., and Cormstock, R. L., "Operation of the Ferrite Junction Circulator," IEEE Trans MTT, January 1965, p. 15.
- Garver, R. V., "Frontiers of Microwave Semiconductor Control Devices," Invited Paper, 1979 International Solid State Devices Conference.
- Garver, R. V., "Microwave Diode Control Devices," Artech House Inc., Dedham, MA, 1976.
- Gaspari, R. A., and Yee, H. H., "Microwave GaAs FET Switching," 1968 IEEE Symposium, MIT Proceedings, p. 49.



Glance, B., "A Fast Low-Loss Microstrip PIN Phase Shifter,"  
Microwave Journal, January 1979, p. 36.

Goel, J., and Wolkstein, H. G., "A 4-8 GHz Dual Gate Amplifier,"  
Int. Sol. State Circuits Conf. Proc., 1978, p. 126.

Helszajn, J., James, D. S., and Wisbet, W. T., "Circulators Using  
Planar Triangular Resonators," IEEE Trans MTT, February 1979, p.  
188.

Hershenov, B., "Microstrip Junction Circulator for Microwave  
Integrated Circuits," IEEE Trans MTT, December 1967, p. 748.

Ito, Y., and Kyogoku, M., "SDMA On-Board Switching System Using  
Rearrangeable Multi-Stage Network," IEEE Int. Conf. Sat. Comm. Syst.  
Tech., April 1975.

Joel, A. E., Jr. "What is Telecommunications Circuit Switching?"  
Proc IEEE, Vol. 65, No. 9, September 1977, p. 1237.

Kane, J., "Fibre Optics Cables Compete with Microwave Relays and  
Coax," Microwave Journal, January 1979, p. 16.

Liechti, C. A., "Microwave Field - Effect Transistors = 1976," IEEE  
Trans MTT, June 1976, p. 279.

Marcus, M. J., "The Theory of Connecting Networks and Their  
Complexity: A Review," Proc IEEE, Vol. 65, No. 9, September 1977,  
p. 1263.

Microwaves, News/International, Vol. 17, No. 11, November 1978, p.  
32.

MIL-HDBK-217B, 20 September 1974, plus Revision 7 September 1976,  
Section 2.2, Table 2.2.2.5, p. 2.2.2-2.

Mumford, W. W., and Scheibe, E. H., "Noise Performance Factors in  
Communications Systems," Horizon House, Dedham, MA, 1968.

Rand McNally Map of Standard Metropolitan Statistical Areas, 1978.

Rosenbaum, F. J., "Integrated Ferrimagnetic Devices, Section IV,"  
Advances in Microwaves 1974, Academic Press, NY.

Salcedo, R. V., and Nations, W. D., "Time Division Multiple Access  
Technology," TRW Report 770331, 1977, Proj. No. 76004921.

Salcedo, R. V., TRW, personal communication 1979.

Siekanowicz, W. W., and Schilling, W. A., "A New Type of Latching Switchable Ferrite Junction Circulator," IEEE Trans MTT, March 1968, p. 177.

Soref, R. A., Sperry Research Center, Sudbury, "Secure Optical Matrix Switch," RADC-TR-79-51. Rome Air Development Center, Air Force Systems Command, Griffiss Air Force Base, April 1979.

Tajima, Y., and Kamihaski, S., "Multiconductor Couplers," IEEE Trans MTT, Vol. 26, No. 10, October 1978, p. 795.

Wu, Y. S., and Rosenbaum, F. J., "Wideband Operation of Microstrip Circulators," IEEE Trans MTT, Vol. 26, No. 10, October 1978, p. 795.

Yamaguchi, M., Ito, Y., and Kyogoku, M., "4 GHz 8 x 8 Switch Matrix for SDMA System," 1975 IEEE-MTTS - International Microwave Conference, Palo Alto, CA.

White, J. F., "Semiconductor Control," Artech House Inc., Dedham, MA, 1977.

Yeh, Y. S., and Reudink, D. O., "The Organization and Synchronization of a Switched Spot-Beam System," ICCS Proceedings 1978, pp. 191-196.

## SECTION 6

### COST PERFORMANCE TRADEOFFS

Critical technologies required to support wideband trunking and direct-to-user services in the 20/30 GHz satellite communications band need to be identified. Previous discussions have pointed out the desirability and advantages of the improved performance and flexibility which result from the use of on-board signal processing.

This section is concerned with the relationship between on-board signal processing and terminal size and complexity. A discussion of satellite and terminal costs and performance/cost tradeoffs is also undertaken, since these issues are central to a determination of technology directions and priorities. Consideration is given to limitations on power flux density as well as conservation of the geostationary arc in setting lower bounds on terminal size.

To properly ascertain technology needs, it is necessary to place such needs within a system context. The system context or exemplar is a point of departure which defines reasonable constraints. It allows ranking of technology needs and provides identification of alternative approaches if a desired technological development fails to occur. The distinction between wideband trunking services and direct-to-user services is important, especially in the tradeoff between terminal and satellite costs. The exemplar system context discussed is presented in figure 6-1.

#### 6.1 GENERAL CONSIDERATIONS

The trunking terminals located in major metropolitan areas will handle only wideband high data rate traffic. This traffic will be concentrated by the terrestrial system and will consist of multiple T3 or T4 trunks. The data rates will vary from a few hundred megabits per second to a few gigabits per second. In the United States, there are 36 metropolitan areas with populations greater than one million persons. Thus, it is expected that the number of trunking terminals will be relatively small. For the purposes of this analysis, a total population of 80 terminals is assumed. This includes dual diversity for mitigation of rain effects. Therefore, 40 trunking centers are considered as the high rate users. This in turn implies that there will be 40 fixed high gain antenna beams on the satellite. Reduction in terminal size and

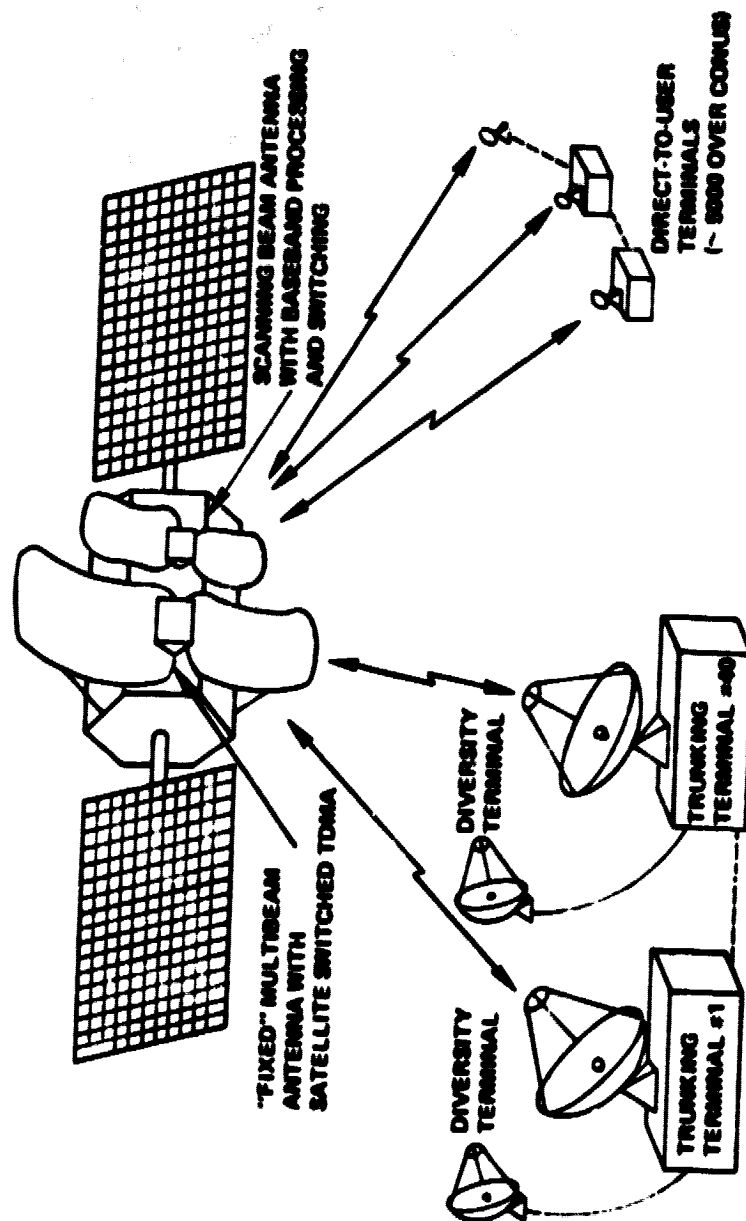


Figure 6-1. Exemplar 20/30 GHz System Concept

complexity is not considered critical in order to minimize system cost and complexity for this trunking service.

The direct-to-user terminals, which are associated with T1 and T2 data rates, will be much greater in number. The total population could be in the tens of thousands, depending upon the cost of the service provided. For the purposes of this analysis and discussion, it is assumed that the population of T1/T2 terminals is on the order of five thousand. Dual diversity at the site is not considered appropriate for these lower cost users. Instead, diversity is provided by the telephone network for those users associated with local telephone switching centers. Where such network diversity is not available, it is assumed that a lower grade of service (i.e., 96.8% rather than 99.9%) is acceptable.

#### 6.1.1 Implications of Power Flux Density Limit

According to the International Telecommunications Union (ITU, 1971), there is a satellite communications allocation in the so-called 20/30 GHz band. The allocation is:

##### Downlink

17.7 - 21.2 GHz

##### Uplink

27.5 - 31.0 GHz

The allocation is for a total bandwidth of 3.5 GHz up and down. In the United States, the upper 1 GHz is reserved for government use. Therefore, for system sizing considerations, 2.5 GHz will be the maximum bandwidth permitted.

The power flux density (pfd) limit is defined in each of the three regions shown in figure 6-2.

##### Region 1 ( $0^\circ < \theta_a < 5^\circ$ )

$pfd < -115 \text{ dBW/m}^2$  in any 1 MHz bandwidth

##### Region 2 ( $5^\circ < \theta_a < 25^\circ$ )

$pfd < -115 + \frac{\theta_a - 5}{2} \text{ dBW/m}^2$  in any 1 MHz bandwidth ( $\theta_a$  is angle of arrival)

##### Region 3 ( $\theta_a > 25^\circ$ )

$pfd < -105 \text{ dBW/m}^2$  in any 1 MHz bandwidth

1A-00,100

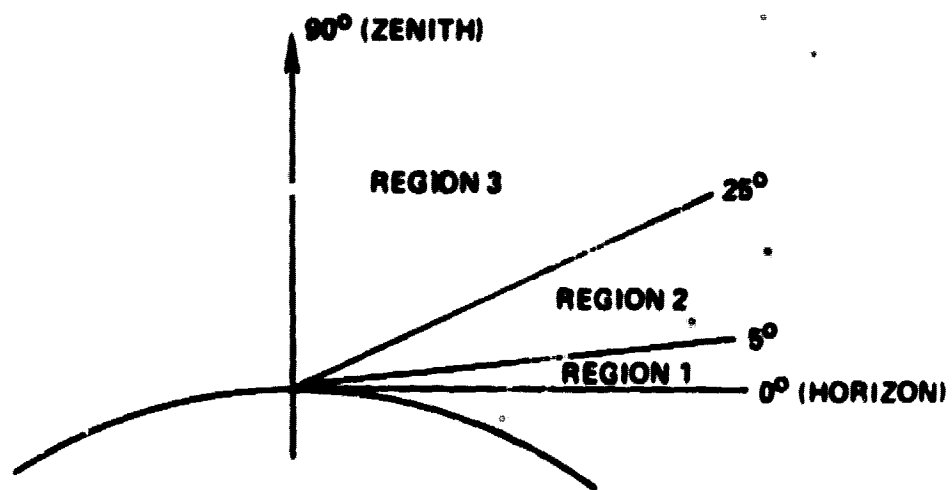


Figure 6-2. Power Flux Density Limit Regions

These limits are for a worst case situation where free space propagation is assumed.

The most stringent pfd limit exists in region 1 and, for purposes of discussion, is the limit chosen. On a per hertz basis, the limit is  $-175 \text{ dBW/m}^2$ . It can be easily shown that this limit translates into an EIRP of  $-11.5 \text{ dBW/Hz}$  for a satellite at geostationary altitude. Figure 6-3 exhibits the relationship between satellite transponder power, antenna diameter, and signal bandwidth for such a geostationary satellite using a nominal downlink frequency of 20 GHz. For example, for a downlink signal bandwidth of 500 MHz, use of a 15-foot diameter satellite antenna implies that a 150 W transmitter is required to achieve an EIRP which reaches the pfd limit. Conversely, use of a 10 W transmitter requires a 55-foot diameter satellite antenna for the same conditions. These are very large satellites, but they would result in the smallest possible terminal.

The implications for smallest possible terminal can be examined now. For region 1, it is easy to show that the terminal quality factor (gain divided by system noise temperature,  $G/T_s$ ) is related to required bit energy per unit noise bandwidth ( $E_b/N_o$ ), pfd ( $\phi$ ), wavelength ( $\lambda$ ), and Boltzmann's constant ( $k$ );

$$\frac{G}{T_s} \geq \frac{E_b}{N_o} \cdot \frac{4\pi k}{\phi \lambda^2} \quad (6.1)$$

which, at 20 GHz, reduces to:

$$\frac{G}{T_s} \geq \frac{E_b}{N_o} - 6.4 \text{ dB/}^\circ\text{K-Hz (log form)} \quad (6.2)$$

In order to size the antenna properly, the system noise temperature,  $T_s$ , must be computed. At 20 GHz, a typical low noise receiver, in the 1985-1990 time frame would have a noise temperature of approximately 50°K for thermo-electrically cooled paramps and 200°K for GaAs FET amplifiers [Cuccia, 1976; Frediani, 1978].

The terminal system noise temperature is given by:

$$T_s = T_{ant} + (L_w - 1) T_o + L_w T \quad (6.3)$$

18-53,101

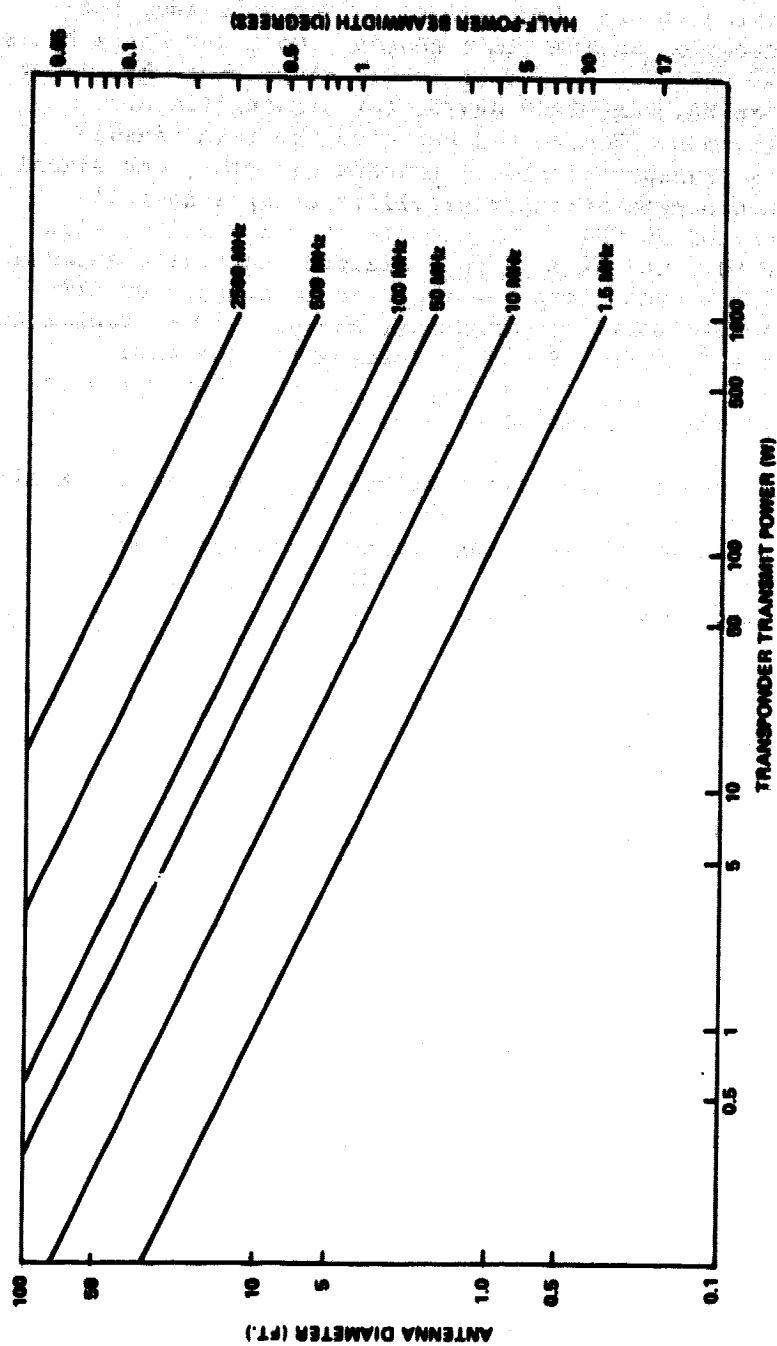


Figure 6-3. Satellite Transponder Power & Antenna Diameter Required to Attain PFD Limit (Region I) as a Function of Bandwidth at 20 GHz (Assumes: 55% Efficiency Antenna, 2.5 dB Feed Loss)



where

$$T_{\text{ant}} = T_{\text{sky}} + T_{\text{tropo}} + T_{\text{atmos}} + T_{\text{rain}} \quad (6.3a)$$

$L_w$  = line loss factor

$T_0$  = room temperature  $\approx 290^\circ\text{K}$

$T$  = first receiver noise temperature (receiver gain assumed sufficiently large to set on overall receiver chain noise figure)

Rain and atmospheric loss are important in determining the minimum antenna size for a system operating at the pfd (or any other EIRP). The factors in equation (6.3a) are elevation angle dependent. With the exception of Alaska and possibly Hawaii (depending upon satellite longitude), elevation angles are generally greater than  $20^\circ$ . In addition,  $T_{\text{sky}}$  and  $T_{\text{tropo}}$  are negligible compared to  $T_{\text{atmos}}$  and  $T_{\text{rain}}$ ; thus, one obtains for a worst case elevation angle of  $20^\circ$ :

$$T_{\text{ant}} \approx T_{\text{atmos}} + T_{\text{rain}} \quad (6.3b)$$

The atmospheric noise temperature is given by:

$$T_{\text{atmos}} \approx T_a (1 - 10^{-A/10}) \quad (6.4)$$

where  $T_a \approx 275^\circ\text{K}$  at 20 GHz and  $A$  is attenuation in dB. At 20 GHz,  $A \approx 1.2$  dB [LeFandre, 1968; Wilder, 1973] which yields,  $T_{\text{atmos}} = 66^\circ\text{K}$ .

$T_{\text{rain}}$  depends upon the desired link availability statistic. It is stated without further justification that an availability of 99.99% is required for trunking terminals and 99.9% for direct-to-user terminals. The system noise temperature,  $T_s$ , for these two availabilities and for the two types of receiver are presented in table 6-1. The figures reflect inclusion of a waveguide and filter loss of 1 dB between the antenna and the low noise amplifier (LNA). In addition, it is also assumed that the LNA has sufficient gain so that it effectively sets the receiver noise figure.

The minimum size antenna values for the trunking and direct-to-user terminals are presented in table 6-2. These values take into

Table 6-1  
Computed System Noise Temperature for  
Indicated Link Availability

Link Availability	Rain Margin (dB)		First Receiver Noise Temperature T <sub>r</sub>	System Noise Temperature, T <sub>s</sub>	
	No Diversity	Dual Diversity		No Diversity	Dual Diversity
99.9%	10	1	{ 50°K 200°K }	494°K 683°K	263°K 452°K
99.99%	30	3	{ 50°K 200°K }	465°K 659°K	349°K 538°K

NOTE:  $G/T$  = antenna gain divided by first receiver noise temperature

$G/T_s$  = antenna gain divided by system noise temperature

account not only the rain margin but also a 3 dB antenna tilt factor. The required  $E_b/N_0$  is taken to be 13.5 dB for the non-regenerative trunking channels and 9 dB for the regenerative direct-to-user channels.

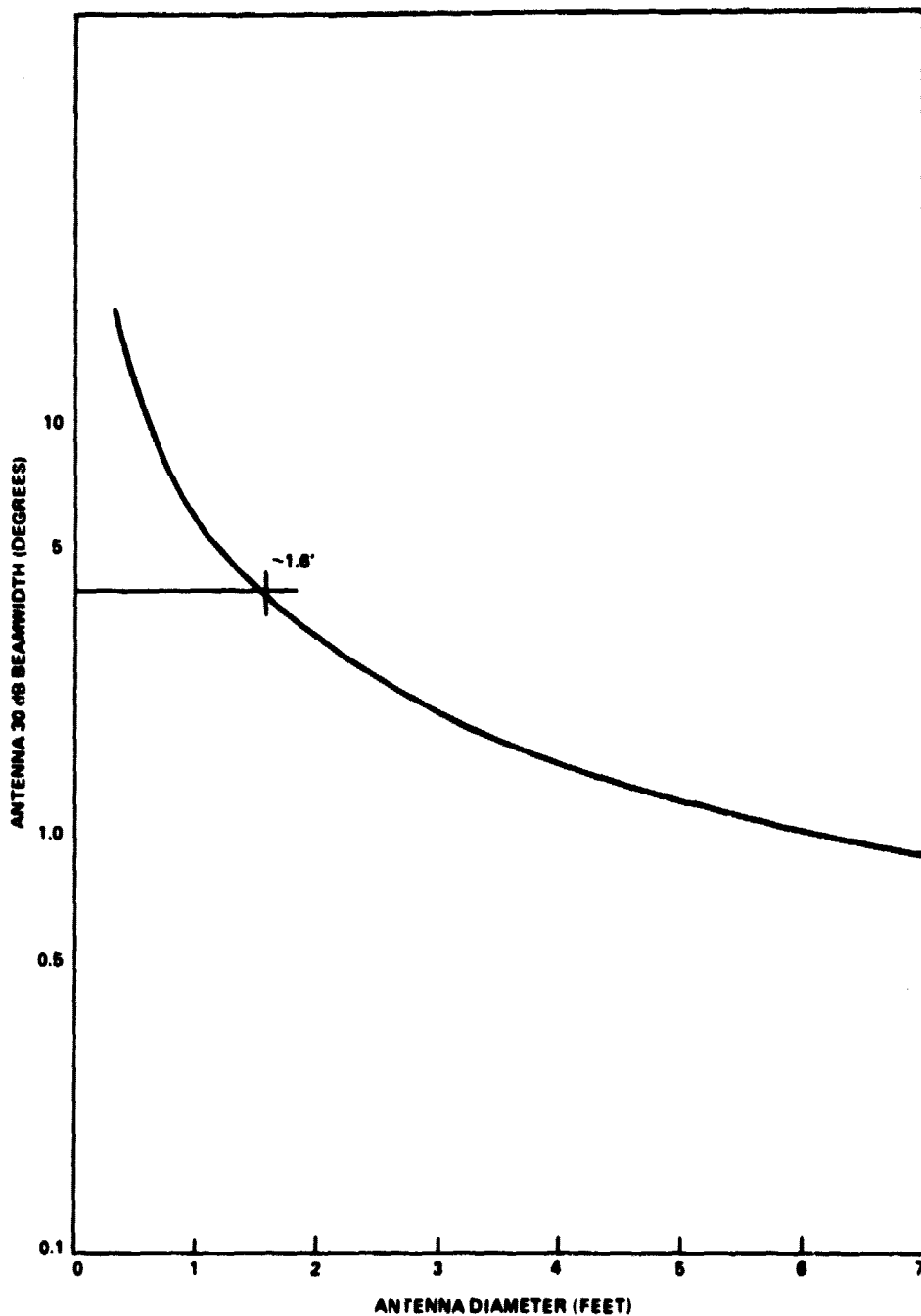
Table 6-2

Minimum Size Antenna  
(at pfd limit)

Link Availability	No Diversity		Dual Diversity	
	T=50°K	T = 200°K	T=50°K	T = 200°K
99.9%	3.1 ft	3.7 ft	0.8 ft	1.1 ft
99.99%	53.0 ft	62.0 ft	2.0 ft	2.5 ft

6.1.2 Geostationary Arc Considerations

The relationship between antenna size and beamwidth is illustrated in figure 6-4 for the uplink frequency band. (The lowest frequency, 27.5 GHz, is chosen.) Since it is important to protect adjacent satellites from being subject to interference from each other's uplinks, it is necessary to select an appropriate angular separation and signal level criterion. Arbitrary, but reasonable, values are selected. These are not final numbers; they merely serve to illustrate a consideration which requires future study and analysis. It is therefore stated without further justification in this analysis that satellites operating in the 20/30 GHz band will not be closer than 4° apart along the geostationary arc. In addition, the signal level permitted to enter an adjacent satellite must be 30 dB below the carrier. This translates into a 30 dB or greater sidelobe criterion. The beamwidth plotted in figure 6-4 is the 30 dB beamwidth. It can be seen from the figure that the smallest permissible aperture at the uplink frequency measures 1.6 feet. Therefore, the satellite EIRP could be reduced for the dual diversity direct-to-user terminal channels or one could accept an additional 3 to 6 dB margin, if operation at the pfd limit were paramount.



1B-55,109

Figure 6-4. Antenna Size vs. 30 dB Beamwidth  
(Assumes: Circular Aperture, 27.5 GHz)

### 6.1.3 Link Imbalance Implications

The 20/30 GHz band is interesting from the point of view of a signal loss imbalance between the uplink and the downlink. This link imbalance is due to several factors: (1) the terminals utilize the same antenna aperture for uplink and downlink, (2) the satellite antennas are separate for transmit and receive but are constant gain (i.e., same spot size for uplink and downlink beams), (3) the frequency differential means different uplink and downlink margins for the same reliability, and (4) there are different system noise temperatures at the satellite and the terminal. The energy per bit per unit noise bandwidth,  $E_b/N_o$  or ENR, required for the uplink and downlink are:

$$ENR_u = EIRP_u - 20 \log \left( \frac{4\pi R}{\lambda_{30}} \right) + \frac{G}{T_s} \Big|_{sat} - k - B - M_u \quad (6.5a)$$

$$ENR_d = EIRP_d - 20 \log \left( \frac{4\pi R}{\lambda_{20}} \right) + \frac{G}{T_s} \Big|_{term} - k - B - M_d \quad (6.5b)$$

where

$$EIRP_u = G_{term}(30) + P_{term};$$

$$EIRP_d = G_{sat}(20) + P_{sat};$$

$$R = \text{range in meters};$$

$$\lambda_{30} = \text{wavelength at 30 GHz};$$

$$\lambda_{20} = \text{wavelength at 20 GHz};$$

$$k = \text{Boltzmann's constant};$$

$$B = \text{channel bandwidth};$$

$$M_u = \text{uplink margin};$$

$$M_d = \text{downlink margin}.$$

The difference in uplink power and downlink power is given by:

$$P_{\text{term}} - P_{\text{sat}} = (\text{ENR}_u - \text{ENR}_d) + (T_{\text{ssat}} - T_{\text{stern}}) + (M_u - M_d) \quad (6.6)$$

For trunking terminals using a single carrier per transponder TDMA signal structure, the required ENR at the downlink terminal,  $\text{ENR}_T$  is related to the uplink and downlink ENR by:

$$\frac{1}{\text{ENR}_T} = \frac{1}{\text{ENR}_u} + \frac{1}{\text{ENR}_d} \quad (6.7)$$

For dual diversity trunking terminals with an availability of 99.99%, a bit error rate (BER) of  $10^{-4}$ , use of CQPSK, a terminal amplifier backoff (B.O.) of 6 dB, and no backoff in the satellite, with a 1600°K satellite system noise temperature, it can be shown that:

$$P_{\text{term}} - P_{\text{sat}} = 8.8\text{dB, or} \\ P_{\text{term}} = 7.6 P_{\text{sat}} \text{ (with 6 dB B.O.)} \quad (6.8)$$

In the case of dual diversity direct-to-user terminals with an availability of 99.9%, a BER of  $10^{-4}$ , DQPSK up and CQPSK down, no backoff, and a 1600°K satellite system noise temperature:

$$P_{\text{term}} - P_{\text{sat}} = 6.5 \text{ dB, or } P_{\text{term}} = 4.5 P_{\text{sat}} \quad (6.9)$$

The conclusion to be drawn from these two equations is that the uplink power and downlink power are related and should not be specified independently, since no advantage occurs in the link performance by over-specifying one or the other. The relationships presented in equations (6.8) and (6.9) result from a certain set of assumptions about the link which can be changed to suit new requirements. Nevertheless, for a given scenario there will be a relationship and it should be evaluated.

#### **6.1.4 Modulation and Signal Processing**

In the satellite system design, it is desirable to reduce the size and weight of the satellite and the size and power of the terminals since this reduces system cost. For the satellite, it is important to reduce the downlink EIRP for a specified terminal  $G/T_s$ . This implies a reduction in the required downlink ENR. Figures 3-27 through 3-29 in section 3 indicate that a reduction in downlink ENR tends to raise uplink ENR. An appropriate operating point should be selected. For purposes of the examples presented here, the operating point is as selected in section 3.

The relationship between downlink EIRP and uplink power is now examined for the non-regenerative repeater and the regenerative repeater for the three backoff conditions illustrated in figures 3-27, 3-28, 3-29, i.e., B.O. = 0/0 dB; 6/0 dB; and 6/4 dB.

This relationship is presented in tables 6-3 to 6-6 for the trunking and direct-to-user terminal. Table 6-3 illustrates the behavior of the non-regenerative trunking channel. The first column states the terminal and satellite power amplifier backoff state. The second and third columns are the optimal uplink and downlink ENRs, respectively. The fourth column is the terminal transmitter power in terms of the satellite transmitter power for the channel. For example, for the B.O. = 0/0 dB case, the terminal transmitter power would be 12 W if the satellite transmitter power were 1 W. The last column provides the relationship between the required downlink EIRP as a function of terminal  $G/T_s$  and channel bandwidth. Table 6-4 presents the data for the non-regenerative direct-to-user channel. Table 6-5 presents the data for the regenerative direct-to-user channel where the regenerating repeater is a hard decision processor. Table 6-6 presents the data for the regenerative direct-to-user channel where the regenerative repeater is a full demodulation/remodulation processor.

Table 6-3

Non-Regenerative Trunking Channel  
(99.99% Availability)

Term B.O., Sat. B.O. (dB)	$ENR_u$ (dB)	$ENR_d$ (dB)	$P_{term}$ (watts)	$EIRP_d +$ $G/T_s - B$ (dB)
0/0	17.5	14.5	$12.0 \times P_{sat}$	0.4
6/0	14.5	13.5	$7.6 \times P_{sat}$	-0.6
6/4	14.0	13.0	$7.6 \times P_{sat}$	-1.1

Table 6-4

Non-Regenerative Direct-to-User Channel  
(99.9% Availability)

Term B.O., Sat B.O. (dB)	$ENR_u$ (dB)	$ENR_d$ (dB)	$P_{term}$ (watts)	$EIRP_d +$ $G/T_s - B$ (dB)
0/0	17.5	14.5	$8.9 \times P_{sat}$	-1.6
6/0	14.5	13.5	$5.6 \times P_{sat}$	-2.6
6/4	14.0	13.0	$5.6 \times P_{sat}$	-3.1



Table 6-5

**Regenerative Direct-to-User Channel  
Hard Decision Processing  
(99.9% Availability)**

Term B.O., Sat. B.O. (dB)	ENR <sub>u</sub> (dB)	ENR <sub>d</sub> (dB)	P <sub>term</sub> (watts)	EIRP <sub>d</sub> + G/T <sub>s</sub> - B (dB)
0/0	14.5	10.5	11.2 x P <sub>sat</sub>	-5.6
6/0	11.8	10.5	6.0 x P <sub>sat</sub>	-5.6
6/4	11.6	10.5	5.8 x P <sub>sat</sub>	-5.6

Table 6-6

**Regenerative Direct-to-User Channel  
Full Demod/Remod  
(99.9% Availability)**

Term B.O., Sat B.O. (dB)	ENR <sub>u</sub> (dB)	ENR <sub>d</sub> (dB)	P <sub>term</sub> (watts)	EIRP <sub>d</sub> + G/T <sub>s</sub> - B (dB)
0/0	11	9	7.1 x P <sub>sat</sub>	-7.1
6/0	9	9	4.5 x P <sub>sat</sub>	-7.1
6/4	9	9	4.5 x P <sub>sat</sub>	-7.1

The best overall operating point for the non-regenerative trunking channels is not apparent. The choice should minimize the total satellite and terminal segment cost. It is premature to seek a global optimum, so a selection is made on a best engineering estimate. The best compromise is the 6/0 dB backoff operating point in table 6-3. A comparison among the non-regenerative, hard decision, and demod/remod processing repeaters for the direct-to-user channels is made in table 6-7, where the 0/0 dB backoff non-regenerative repeater is used as the baseline. This is accomplished by setting the downlink transmit power at 1 W and comparing all other cases with it. The transmitter powers presented include backoff effect.

Table 6-7

## Direct-to-User Processor Comparison

Term B.O.						
Sat B.O.	Non-Regenerative		Hard Decision		Demod/Remod	
(dB)	P <sub>sat</sub> (W)	P <sub>term</sub> (W)	P <sub>sat</sub> (W)	P <sub>term</sub> (W)	P <sub>sat</sub> (W)	P <sub>term</sub> (W)
0/0	1	8.9	0.4	4.5	0.3	2
6/0	0.8	17.7	0.4	9.6	0.3	5
6/4	1.8	16.0	1.0	9.6	0.7	5

The full demod/remod case with 0/0 dB backoff yields the smallest satellite and terminal RF transmitter size. Consequently, the recommended satellite processor configuration is: (1) a satellite switched TDMA non-regenerative processor with 6/0 dB backoff for the fixed beam trunking channels, and (2) a full demod/remod processor with 0/0 dB backoff with baseband switching for the scanning beam direct-to-user channels. It is too soon to evaluate whether the burst rate on the direct-to-user channels should be equal on the uplink and downlink. It appears reasonable to expect that FDMA on the uplink with a TDM downlink will be lower cost than TDMA up and TDM down, due to reduced terminal size and complexity. This issue is examined more carefully in the system cost/performance tradeoffs discussion.

## 6.2 ACCESS METHODS

Before terminal and satellite design implications of on-board processing are discussed, the channel bandwidths appropriate for trunking and direct-to-user terminals are examined briefly. The modulations considered provide approximately 1.4 information bits per channel symbol. The total allocation for commercial use at 20/30 GHz within CONUS is 2.5 GHz. For purposes of discussion, 100 MHz appears useful for the direct-to-user terminals and 2400 MHz for the trunking terminals. It is further suggested that the 2400 MHz be divided into four 600 MHz wide contiguous bands. There would be four TDMA carriers in the 2400 MHz band but each carrier would access its own 600 MHz wide transponder. Thus, in the satellite there would be only a single carrier per transponder and four transponders per beam. This would provide a degree of redundancy in the satellite as well as reducing the terminal modem complexity and

cost. Each 600 MHz band could carry three T4 trunks or eighteen T3 trunks, or some combination of the two.

The choice of 100 MHz for the direct-to-user terminals was made for two reasons: first, to reduce terminal RF and modem costs, and second, to provide an uplink data rate for which it is reasonable to expect on-board processing capabilities in the 1985-1990 time frame. For the direct-to-user channel, two uplink access strategies are considered.

#### 6.2.1 Case 1. TDMA

The equal burst rate approach implies that there are 100 simultaneous T1 users, each bursting at approximately 154 Mb/s on the uplink, with a composite 154 Mb/s downlink data rate. The band occupancy is approximately 100 MHz. Although the feasibility of on-board baseband processing at these rates is uncertain, they do provide an interesting upper bound and a goal for on-board demodulation rates. The total number of simultaneous T1 users in CONUS depends on the number of beams. Two scanning beams could probably service 5000 terminals with a reasonable grade of service.

#### 6.2.2 Case 2. FDMA

The unequal burst rate approach suggests ten uplink carriers per access with a burst rate of approximately 15 Mb/s per carrier. On-board processing would make it feasible to restack these users onto a 154 Mb/s composite downlink. Again, 100 simultaneous T1 users can be accommodated, albeit with slightly less flexibility and efficiency since 10 users must be in each beam access, and the scan rate would be one tenth that in case (1). Again, the downlink bandwidth would be 100 MHz but the total uplink bandwidth of 100 MHz would be divided into ten 10 MHz bands.

The relationship between terminal  $G/T_0$ , satellite antenna size, and satellite RF power for the trunking terminals and the direct-to-user terminals is now examined. It can be recalled from tables 6-3 and 6-6 that:

$$\text{Trunking Channel: } P_{\text{sat}} + G_{\text{sat}} = 87.2 - G/T_0 \quad (6.10a)$$

$$\text{Direct-to-User Channel: } P_{\text{sat}} + G_{\text{sat}} = 72.9 - G/T_0 \quad (6.10b)$$

The results of equations (6.10a) and (6.10b) are plotted in figures (6-5a) and (6-5b), respectively. Included in the figures is a range of power over which each of three types of 20 GHz power amplifier suitable for use on the satellite is considered. Power levels less than 5 W are suitable for solid state amplifiers in the 1985-1990 time frame, although development will be required for the greater than 1 W power levels. Helix TWT technology is suitable for the 1 W to 20 W range, although some development would be required for power levels greater than 10 W. Coupled cavity TWAs are suitable for the higher power levels. Development would be required at power levels greater than 20 W.

From the point of view of RF amplifier reliability and weight in the satellite, it is clear from figure 6-5a that the trunking channels would benefit from the use of higher gain antennas ( $D > 15$  ft) on the satellite and higher G/T terminals ( $G/T > 25$  dB - °K). This would permit use of solid state power amplifiers. In the case of direct-to-user channels (figure 6-5b), use of solid state power amplifiers would appear to be the solution since larger satellite antennas and G/T terminals are not required for that power range.

The lower power solid state power amplifiers are desirable in the terminal segment as well, for reasons of cost and reliability. Since it has been shown that the uplink power for trunking terminals is 7.6 times the satellite downlink power (not including terminal backoff), lower satellite powers are indicated. An analogous argument holds for the direct-to-user channels.

The best balance among satellite RF power, satellite antenna size, and terminal G/T for a best system design is not obvious. In order to best determine what critical spacecraft technologies require development, not only in the on-board processor elements but also in the RF area, an analysis is needed. This process should involve the terminal segment as well. The optimum solution should yield the minimum cost system for a constant performance level.

### 6.3 SATELLITE WEIGHT OPTIMIZATION

A balance among satellite antenna size, RF power, and terminal G/T is desirable. This implies a minimum system cost for constant performance. In order to minimize system cost, there must be a tradeoff between satellite cost and terminal segment cost. Satellite cost is proportional primarily to satellite weight which in turn depends upon satellite EIRP and other factors. In order to minimize satellite cost, one would minimize satellite EIRP. However, a decrease in satellite EIRP is accompanied by an increase

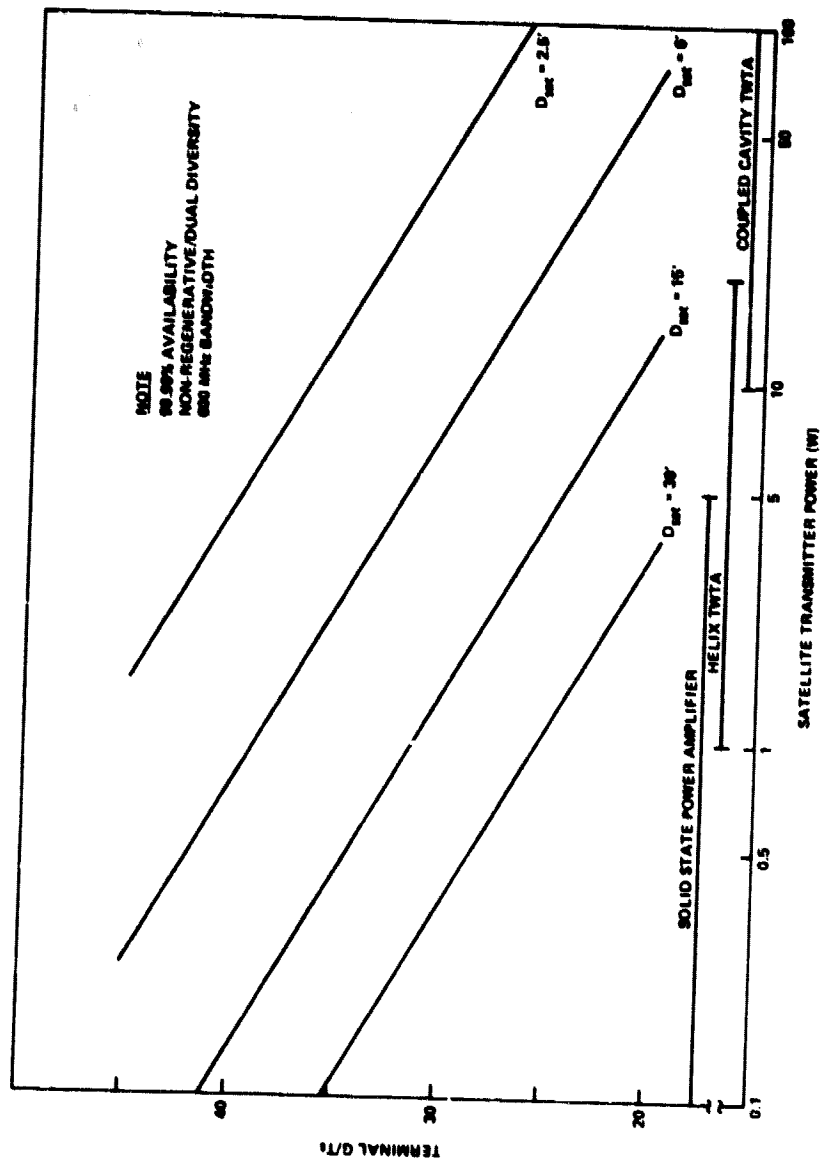


Figure 6-5a. Satellite Transmit Power vs. Trunking Terminal G/Ts

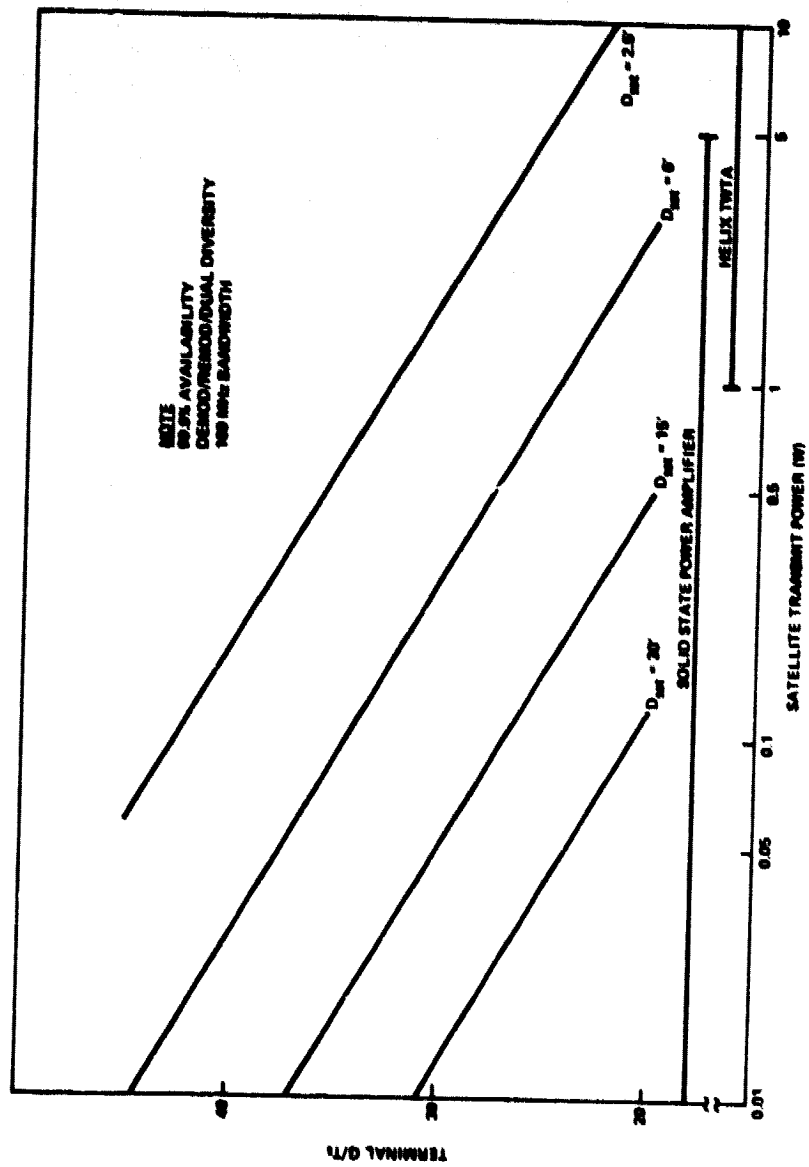


Figure 6-5b. Satellite Transmitted Power vs. Direct-to-User Terminal G/Ts

in terminal G/T so as to maintain constant performance. This increases terminal segment cost. Therefore, a global minimization is necessary.

The approach taken in this first examination of the problem is to determine the minimum weight satellite as a function of EIRP. In order to accomplish this, a model is suggested. It should be recognized that the results which derive from the model will change if the model is changed. The major purpose of the analysis presented is to demonstrate that selection of technologies for development must be based on reasonable and defensible system needs. By following this approach, the technology developments required can be prioritized in a rational manner and alternative technologies and fall back positions can be readily identified.

The satellite model is shown in figure 6-6. For a specific system performance level and on-board processing selection, the only variables which significantly affect satellite weight are the antennas and the RF power amplifiers. The weight of the satellite antennas, including the feeds and transmitters (and the weight due to thermal power control and primary electrical power), can be modeled.

According to a Bell Telephone Laboratory report [BTL, 1968] performed for NASA, a rigid reflector of the type required for use at 20/30 GHz, has a weight proportional to antenna diameter, i.e.,  $W_t = 10D$ , where D is in feet. For the purposes of this analysis, the antennas being modeled include the reflector, feeds, and structure. For an N-beam antenna, the weight model is:

Item	Weight (lb)
reflector	10D
feed	N/10
structure	15% (10D + N/10)

Therefore,

$$W_{ant} \approx 12D + N/10 \text{ (lbs)}, \quad D = \text{diameter in feet}, \quad (6.11)$$

$N = \text{number of feeds}$

4-20-78

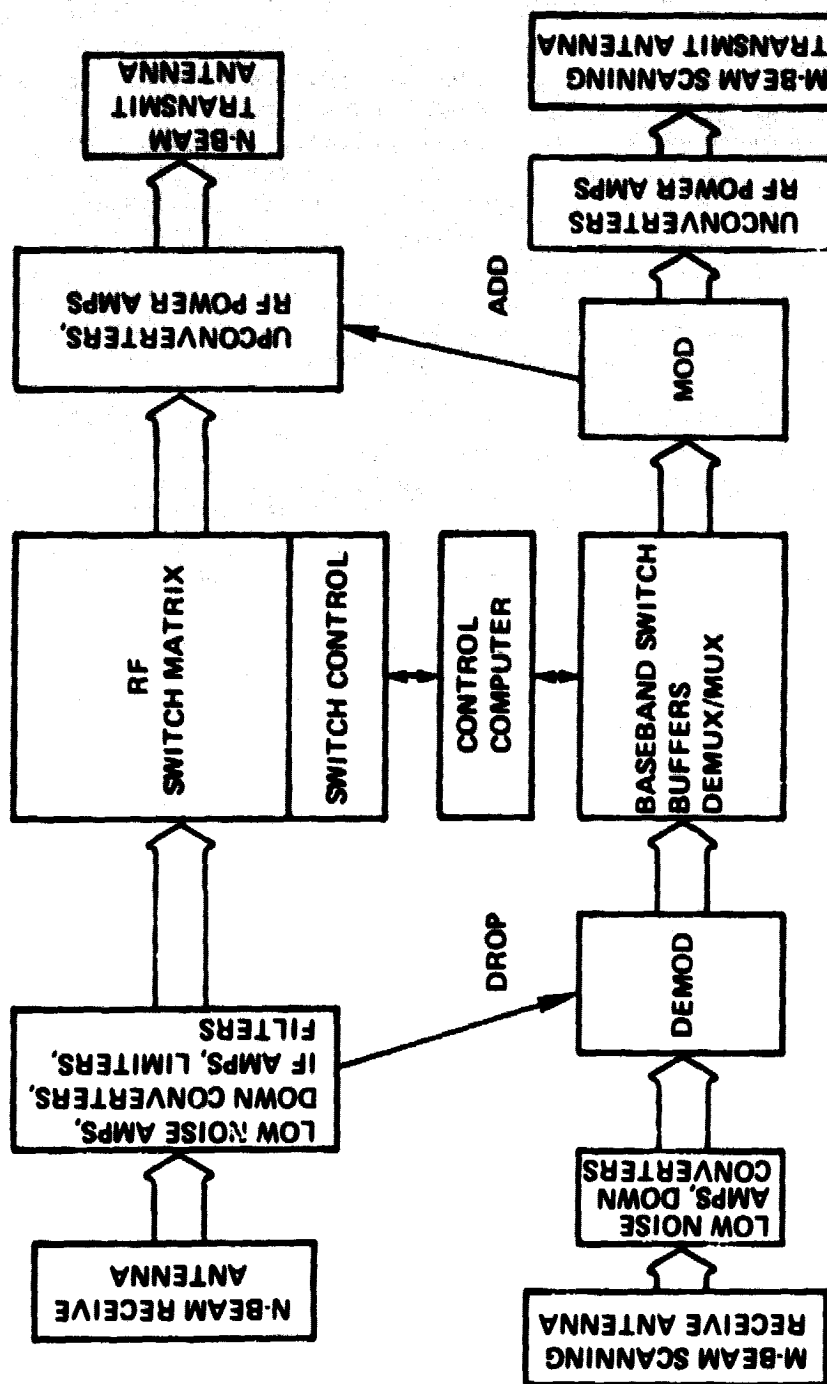


Figure 6-6. Satellite Model



In the case of the downlink transmitters and the weight burden due to the primary electrical power requirements, available data indicate:

TWTA: for power levels greater than 1 W

$$W_{\text{TWTA}} \approx 6 + p/3 \text{ (lb); } p = \text{RF power in watts} \quad (6.12)$$

Solid State PA: for power levels less than 5 W

$$W_{\text{SSPA}} \approx p \text{ (lb)} \quad (6.13)$$

Prime Power Weight Burden:

$$W_{\text{pp}} \approx 0.4 p/\eta \text{ (lb); } \eta = \text{DC to RF efficiency} \quad (6.14)$$

Thus, for a TWTA efficiency of 25% and a projected solid state amplifier efficiency of 25%, one obtains the following approximations for overall antenna and transmitter weight:

(a)  $P > 1 \text{ W}$ , TWTAs

$$W_t \approx 12D + 6N + 2.5Np \text{ (lb)} \quad (6.15a)$$

(b)  $P < 5 \text{ W}$ , solid state amplifiers

$$W_t \approx 12D + 0.1N + 2.5Np \text{ (lb)} \quad (6.15b)$$

One can then minimize the weight due to transmitters and antennas by holding EIRP (or performance) constant on a per beam basis. The EIRP is related to power and gain on a per beam basis in the following manner:

$$\text{EIRP} = pG,$$

where

$$p = \text{transmit power/beam in watts} \quad (6.16)$$

$$G = \text{antenna gain for a beam}$$

Since,

$$G = CD^2,$$

$$D = \sqrt{\frac{EIRP}{C_p}} ; C \approx 1260 \text{ at } 20 \text{ GHz} \quad (6.17)$$

Using this relationship, one can minimize  $W_t$ . For either the TWTA or solid state power amplifier case, the minimum weight antenna/transmitter package for an N-beam system is defined when the RF transmit power and antenna diameter are:

$$P_{\min} = \left[ \frac{\sqrt{EIRP}}{14.8N} \right]^{2/3} \quad \text{and} \quad D_{\min} = \sqrt{\frac{EIRP}{C_{p_{\min}}}} \quad (6.18a,b)$$

Consequently, one can obtain, for the model chosen here, a minimum weight configuration for a specified satellite EIRP. Before examining these results in more detail, it is necessary to further define a strawman system upon which to base various tradeoffs. The total range of tradeoffs is limited in this report. However, the approach and methodology are sufficient to cover a wider range of alternatives during the second phase of this effort.

#### 6.4 TERMINAL AND SATELLITE TRANSPONDER DESCRIPTIONS

Previous discussions of modulation, data rates, bandwidths, and terminal quantities have indicated that there are two distinct services being provided, i.e., trunking and direct-to-user. This section does not discuss the drop and add features which provide the interconnection between the users. For the bulk of traffic, these services are provided by two separate and distinct transponders on the same bus, as indicated in figure 6-6. In the strawman system, a separate antenna system as well as separate RF subsystems and baseband subsystems are provided for each channel type. A specific definition for each channel type follows without further justification. This strawman is used to illustrate the cost/performance tradeoffs and the implications for technology development.

#### **6.4.1 Trunking Channel Description**

Number of beams	- 40
Number of terminals	- 80 (dual diversity included)
Channel availability	- 99.99%
Modulation	- CQPSK (non-regeneration)
Access method	- TDMA
Channel bandwidth	- 600 MHz
Number of channels/beam	- 4
Satellite/terminal HPA bandwidth	- 2400 MHz
Terminal G/Ts	- 20 dB-°K to 36.1 dB-°K
Total data rate	- 3.4 Gb/s

Block diagrams of the satellite transponder and the trunking terminal are presented in figures 6-7 and 6-8, respectively.

#### **6.4.2 Direct-to-User Channel Description**

For the direct-to-user channel, a system configuration which consists of two simultaneous uplink beams and two simultaneous downlink beams has been chosen. This transponder is capable of supporting 200 simultaneous T1 users, which is sufficient to service a population of 2000 to 5000 terminals. There are three processing options available for the scanning beam direct-to-user channel: no regeneration with TDMA, regeneration with TDMA up/TDM down, and regeneration with FDMA up/TDM down.

##### **● Non-regenerating Direct-to-User Channel**

Number of beams	- 2
Number of terminals	- 5000
Channel availability	- 99.9%
Modulation	- CQPSK
Access method	- TDMA
Channel bandwidth	- 100 MHz
Number of channels/beam	- 1
Satellite/Terminal HPA bandwidth	- 100 MHz
Terminal G/Ts	- 11.8 dB-°K to 30.8 dB-°K
Total data rate	- 300 Mb/s

Block diagrams of the transponder and terminal are shown in figures 6-9 and 6-10, respectively.

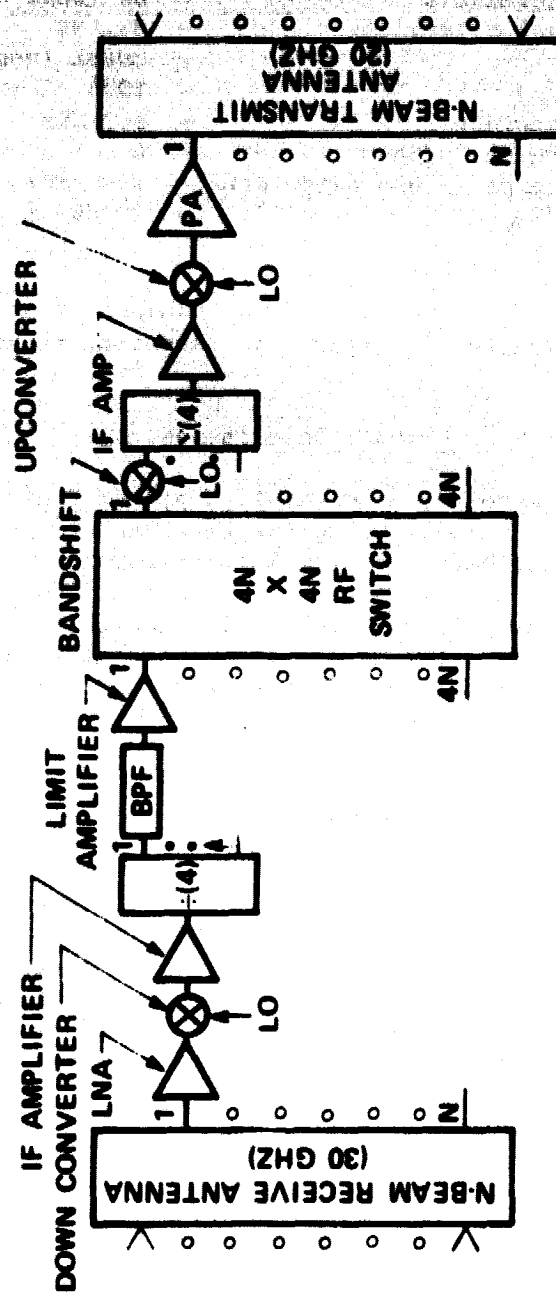


Figure 6-7. Trunking Transponder

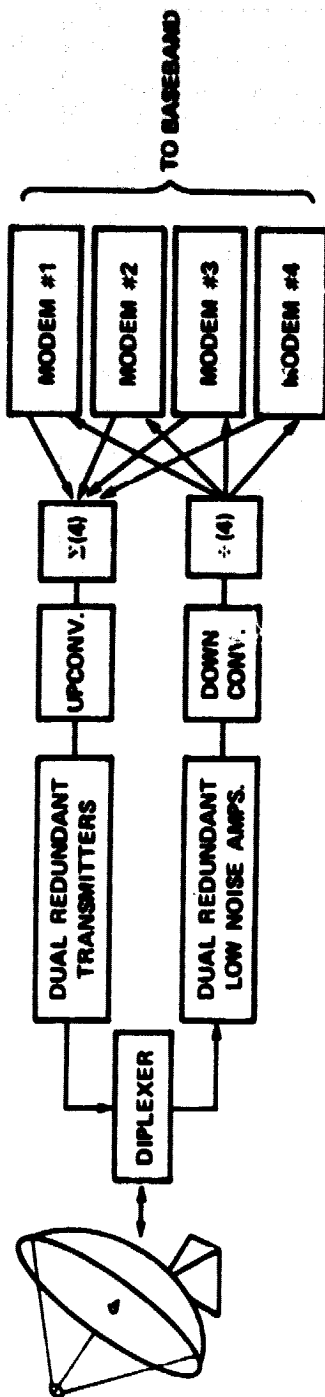


Figure 6-8. Trunking Terminal

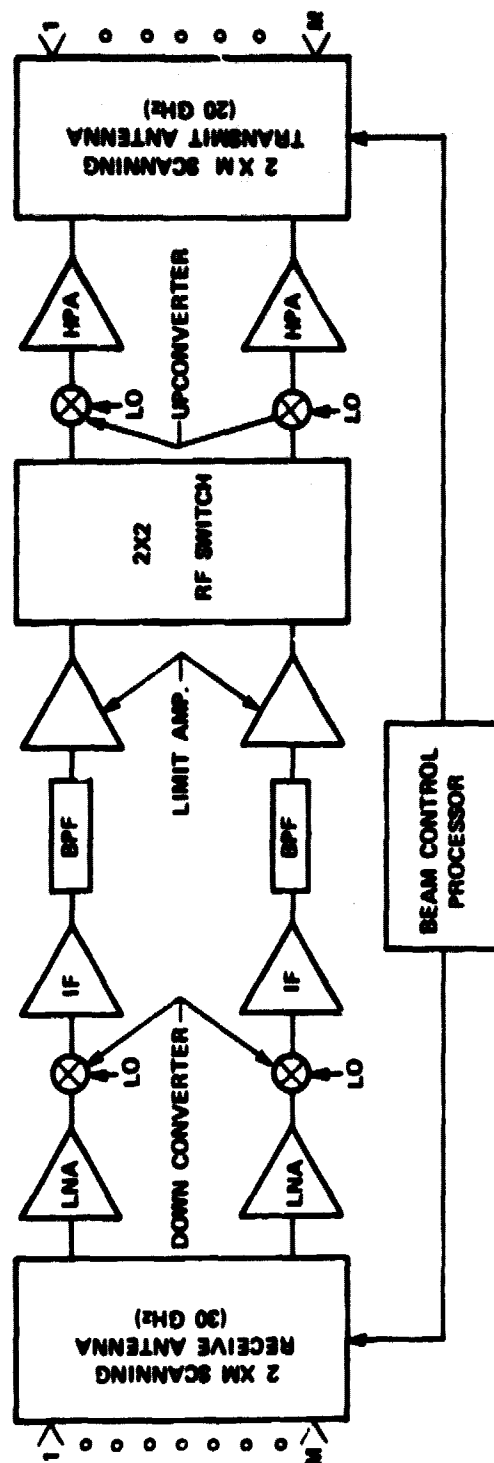


Figure 6-9. Non-Regenerative Direct-to-User Transponder

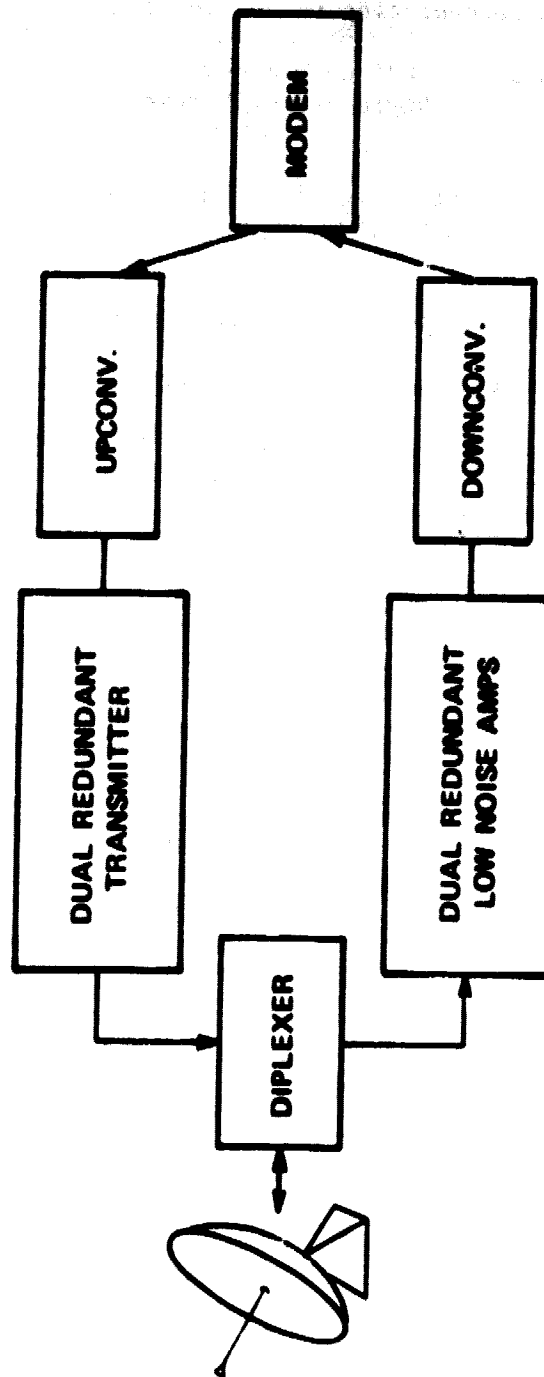


Figure 6-10. Direct-to-User Terminal

● Regenerating TDMA Direct-to-User Channel

Same as non-regenerating except that:

Access Method - TDMA up/TDM down  
Modulation - DQPSK up/CQPSK down

Block diagrams of the transponder and terminal are shown in figures 6-11 and 6-10, respectively.

● Regenerating FDMA Direct-to-User Channel

Same as non-regenerating except that:

Modulation - DQPSK up/CQPSK down  
Access method - FDMA up/TDM down  
Channel bandwidth - 10 MHz per FDMA carrier up/100 MHz down  
Number FDMA channels - 10

Block diagrams of the transponder and terminal are shown in figures 6-12 and 6-10, respectively. It should be noted that MSK would probably be used on the uplink to reduce crosstalk on the multicarrier uplink channel. However, the link performance would be approximately that of DQPSK.

## 6.5 SATELLITE AND TERMINAL COST ESTIMATION

In order to trade off alternative approaches, the cost of both the space segment (satellite) and the terminal segment must be examined. The cost figures presented in this section are valid for this comparison only.

Satellite cost estimation can be based upon the very detailed SAMSU Unmanned Spacecraft Cost Estimation Model or on a simpler rule-of-thumb basis which relates the cost of the spacecraft to its total weight. The latter approach is selected as most appropriate for this report. It has been shown (MILSATCOM, 1976) that:

Recurring Cost (FY 79 \$)

$$C_R (\$M) \approx 0.031 (\text{on-orbit weight, lb})^{0.93} + \text{launch} \quad (6.19a)$$



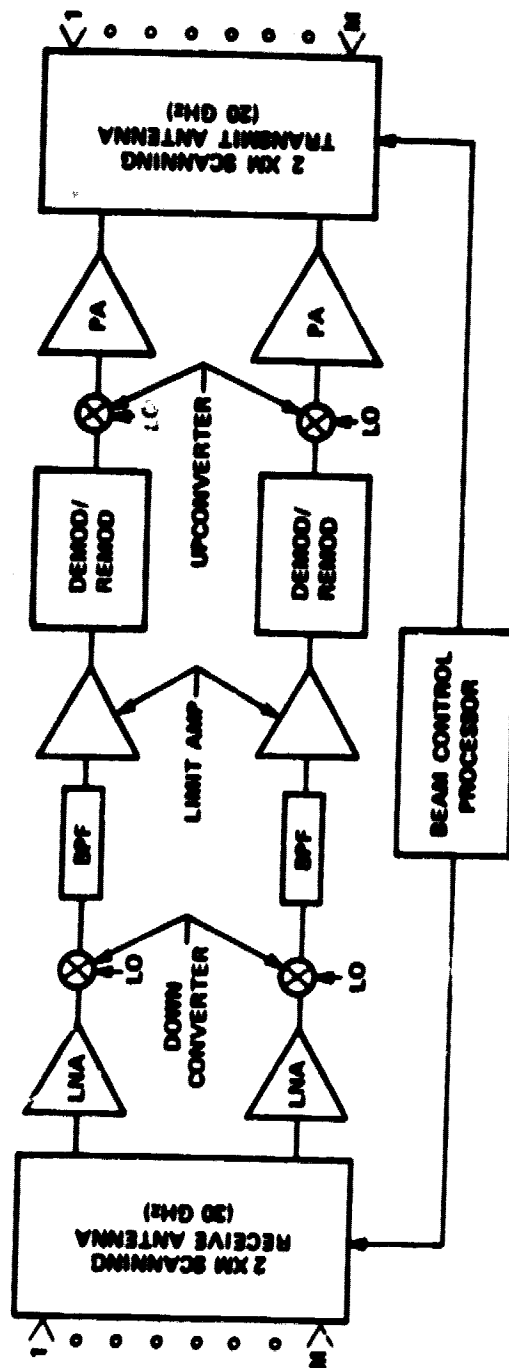


Figure 6-11. Regenerative TDMA Direct-to-User Transponder

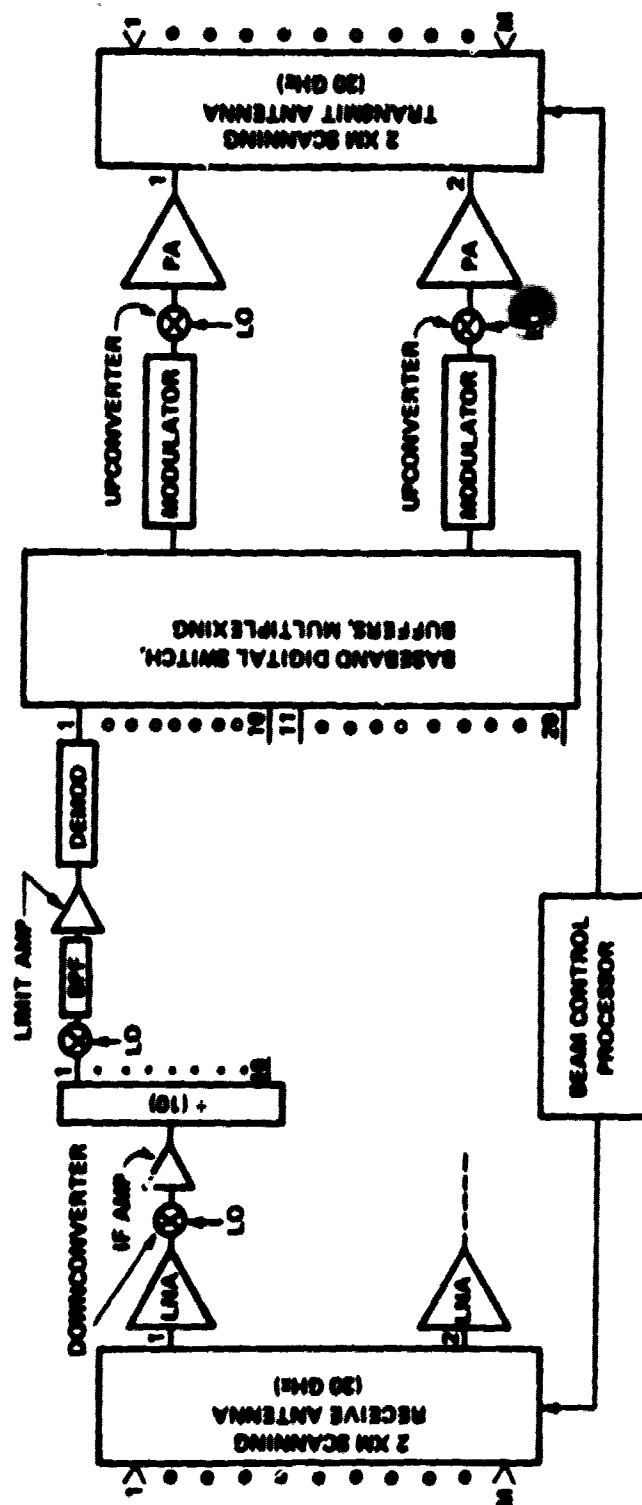


Figure 6-12. Regenerative FDMA Direct-to-User Transponder

Non-recurring (development) cost (FY 79 \$)

$$C_{NRE} (\$M) \approx 0.016 (\text{on-orbit weight, lb})^{1.15} \quad (6.19b)$$

The launch cost is ignored in this report. It can be shown that the on-orbit weight of a communications satellite is directly related to the weight and power requirements of the communications package.

As approximation for the beginning of mission (BOM) weight (lb) and power (watts) for a communications satellite is:

$$W_t \approx 800 + 1.53w_c + 0.8 p_c \quad (6.20a)$$

$$P_t \approx 240 + 1.26 p_c \quad (6.20b)$$

where,  $w_c$  is the weight of the communications package and  $p_c$  is the d.c. power required by the communications package.

The weight and power requirements for the various transponder packages described previously are:

● Non-regenerating Fixed Beam Transponder

$$w_c(a) \approx 33 + 12[D(20) + D(30)] + 16.8N + 1.3Np + 0.3N^2 \quad (6.21a)$$

$$p_c(a) \approx 35 + 12.2N + 9Np + 0.3N^2 \quad (6.21b)$$

where

$D(20)$  = diameter of receive antenna in feet

$D(30)$  = diameter of transmit antennas in feet

$N$  = number of fixed beams

$p$  = satellite power amplifier power

For  $N = 40$ :

$$w_c(a) \approx 1200 + 12[D(20) + D(30)] + 50p, \text{ and} \quad (6.22a)$$

$$p_c(a) \approx 1000 + 360p \quad (6.22b)$$

● Non-regenerating Direct-to-User Transponder (2 beams)

$$w_c(b.1) \approx 40 + 12[D(20) + D(30)] + D^2(20) + 2p \quad (6.23a)$$

$$p_c(b.1) \approx 40 + 2D(20) + 10p \quad (6.23b)$$

● Regenerating TDMA Direct-to-User Transponder (2 beams)

$$w_c(b.2) \approx 225 + 12[D(20) + D(30)] + D^2(20) + 2p \quad (6.24a)$$

$$p_c(b.2) \approx 430 + 2D(20) + 10p \quad (6.24b)$$

● Regenerating FDMA Direct-to-User Transponder (2 beams)

$$w_c(b.3) \approx 415 + 12[D(20) + D(30)] + D^2(20) + 2p \quad (6.25a)$$

$$p_c(b.3) \approx 720 + 2D(20) + 10p \quad (6.25b)$$

Using the above equations with the appropriate expressions for  $ERP_d$  from tables 6-3 and 6-6 and equations (6.18a) and (6.18b) yields the communications weight and power estimates. These are then used in equation (6.20a) to obtain the on-orbit satellite weight for each case. Recurring and non-recurring cost can then be obtained using equations (6.19a) and (6.19b).

Cost estimates for the trunking and direct-to-user terminals are now developed. The following cost estimating relationships were used for the antennas.

For antennas less than 15 feet in diameter, fixed pointing is assumed, and

$$D < 15': C_{ant} (\$K) \approx 0.12D^2, \text{ } D \text{ is antenna diameter} \quad (6.26a)$$

in feet

For antennas greater than 15 feet in diameter, auto tracking is assumed to be required, and

$$D > 15': C_{ant} (\$K) \approx 20 + 1.18D^{3/2} \quad (6.26b)$$

Two types of LNA are considered appropriate for use: a thermoelectrically cooled parametric amplifier with noise temperature of 50°K and a GaAs FET amplifier with noise temperature of 170°K. Both of these performance levels are anticipated to be available in the 1985 time frame. The estimated costs of these LNAs in 1979 dollars are:

$$C_{param} (\$K) \approx \$100K \quad (6.27)$$

$$C_{FET} (\$K) \approx \$10K \quad (6.28)$$

The cost of an antenna with a dual redundant LNA system is shown as a function of terminal G/T in figure 6-13. It should be noted that terminal G/T does not equal  $G/T_g$  since the effect of antenna noise and rain noise are not included. An equivalence between G/T and  $G/T_g$  for trunking terminals and direct-to-user terminals is shown below in tables 6-8 and 6-9. The lowest cost configuration is chosen, which means that GaAs FET LNAs are used when  $G/T \gtrsim 45 \text{ dB-}^\circ\text{K}$ .

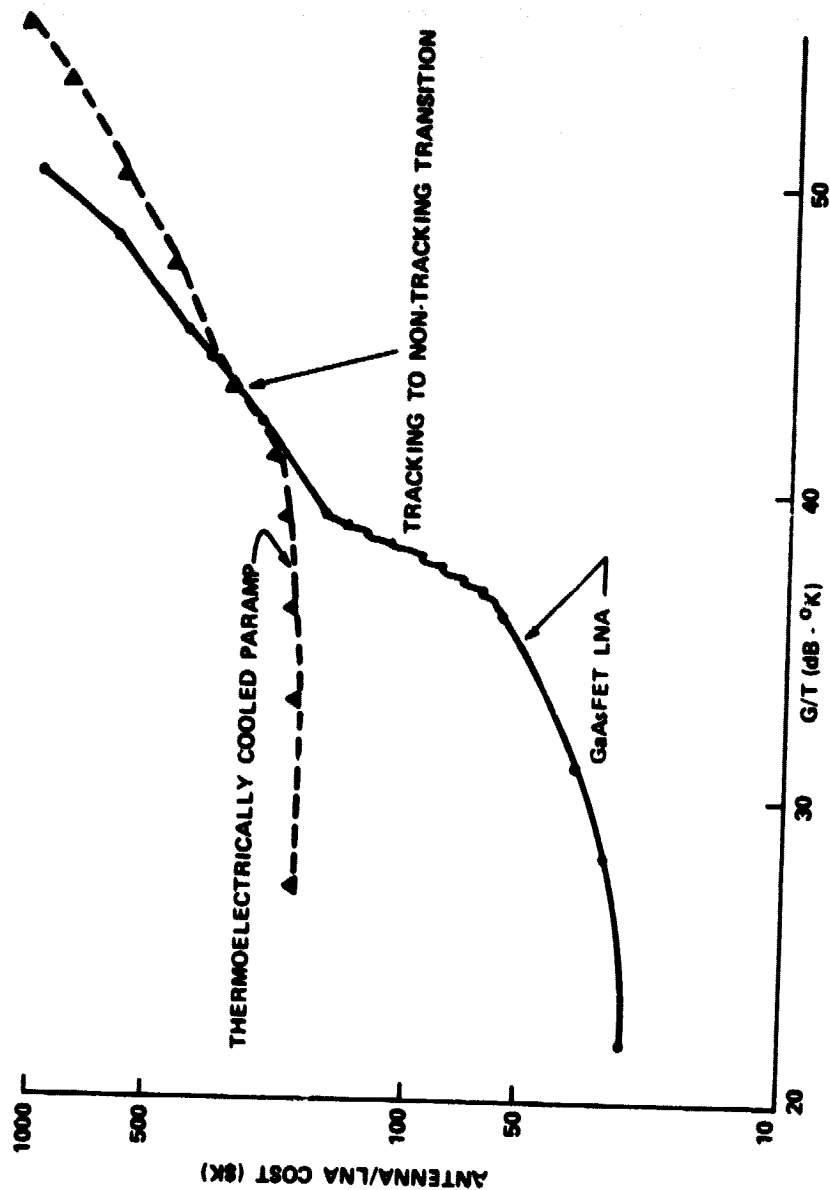


Figure 6-13. Antenna/LNA Cost vs. G/T

Table 6-8

Trunking Terminal G/T and G/T<sub>s</sub>

G/T (dB-°K)	G (dB)	G/T <sub>s</sub> (dB-°K)
25.2	48.3	20.0
33.1	56.1	27.8
36.6	59.8	31.5
40.0	63.0	34.8
43.0	66.0	37.7
46.0	63.0	36.1

Table 6-9

Direct-to-User G/T and G/T<sub>s</sub>

G/T (dB-°K)	G (dB)	G/T <sub>s</sub> (dB-°K)
17.0	40.0	11.8
20.0	43.0	14.8
23.0	46.0	17.6
26.0	49.0	20.8
29.5	52.5	24.3
33.0	56.0	27.8
36.0	59.0	30.8

The cost of the transmitter as a function of power is shown in figure 6-14. This figure assumes that only TWIAs are used. At the lower power levels, the transmitter cost is dominated by the high voltage power supply. In the 1985 time frame, an all solid state transmitter in the 5 to 10 W range may be feasible but it is not clear that it would be less expensive.

The cost of the up/down converters is estimated to be approximately \$10K. The small quantity cost of the RF portion of the terminal is given by the following:

$$C_{RF}(\$K) \approx 1.7 [C_{ant} + C_{LNA} + C_{trans} + C_{conv}] \quad (6.29)$$

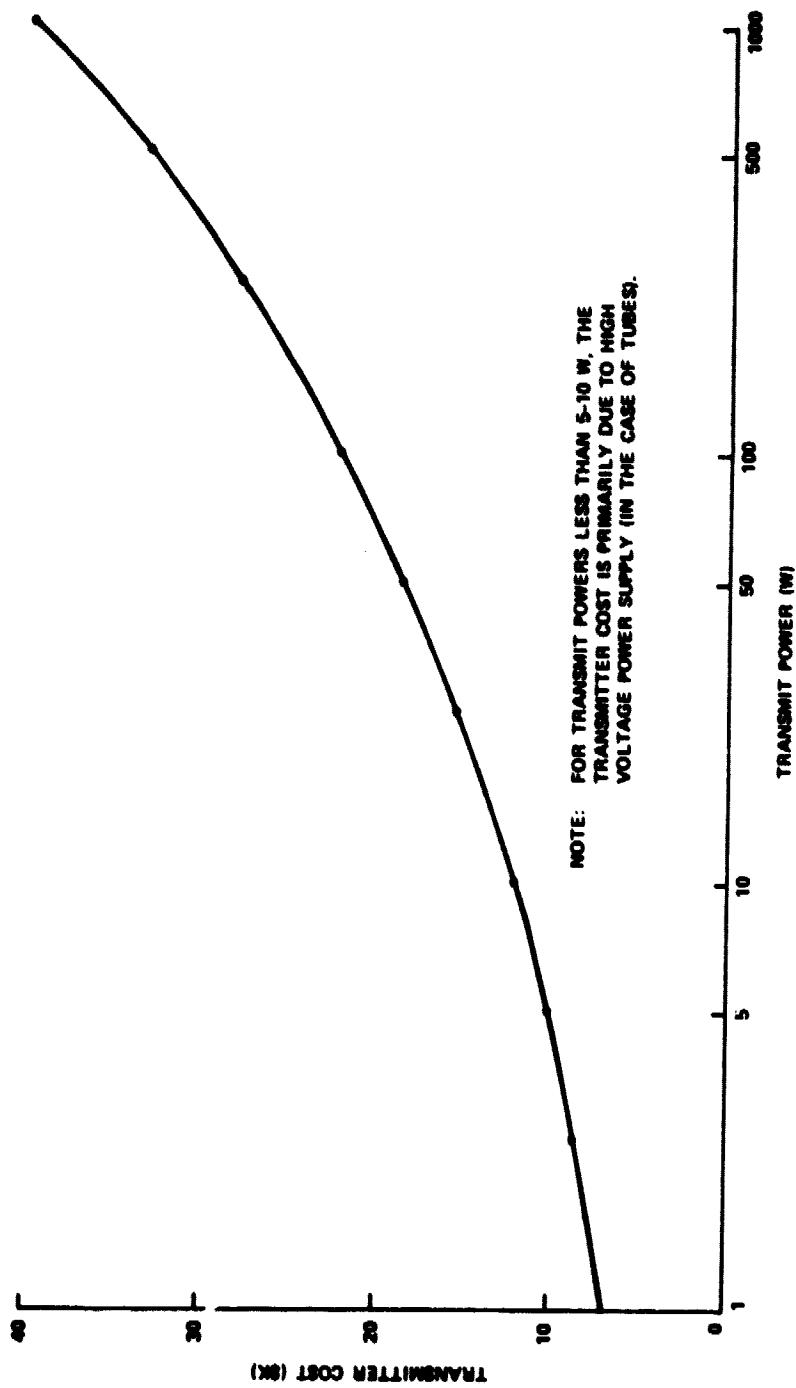


Figure 6-14. Transmitter Cost vs. Transmitter Power (30 GHz)



The cost of a high rate TDMA trunking modem, with a data rate on the order of 800 Mb/s, is estimated to be \$400K in small quantities. The cost of the high rate direct-to-user TDMA modem, with a data rate of 150 Mb/s, is estimated to be \$160K in small quantities. The cost of the lower rate direct-to-user FDMA modem, with a data rate of about 15 Mb/s is estimated to be approximately \$100K in small quantities. For convenience, a single learning curve factor of 0.95 is used for estimating the cost of a large number of terminals. The unit costs as a function of quantity (Q) for this factor are listed in table 6-10.

Table 6-10

Relative Unit Costs  
(Quantity Effect)

<u>Q</u>	<u>Relative Cost</u>
1	100%
100	70%
1000	60%
5000	50%

The cost of a complete terminal (for  $Q \sim 1$ ) is given by:

$$C_{\text{term}} (\$K) \approx C_{\text{RF}} + C_{\text{modem}} \quad (6.30)$$

## 6.6 EXEMPLAR SYSTEM TRADEOFFS

The cost/performance tradeoffs for three exemplar systems are examined in order to determine what the optimum size satellite and terminal configurations should be. Terminal costs are so dominant that there is a great incentive to expend resources to improve satellite performance. The three exemplar systems are: (1) non-regenerative trunking transponder plus non-regenerative TDMA direct-to-user transponder; (2) non-regenerative trunking transponder plus regenerative TDMA direct-to-user transponder; and (3) non-regenerative trunking transponder plus regenerative FDMA direct-to-user transponder. For this first cut, the direct-to-user and trunking transponders use separate antennas and feed structures. The

transponder weights and powers for trunking and direct-to-user channels are computed as a function of terminal  $G/T_g$ . These are used to compute satellite weight and cost. In addition, terminal segment cost is computed for each terminal  $G/T_g$  and a total system cost is determined. The minimum cost configuration is used to determine preliminary technology development priorities.

The non-regenerative trunking transponder which is common in all three exemplars is examined first. The satellite implications are presented in table 6-11. The terminal segment and total cost attributable to the trunking system are presented in table 6-12. A typical trunking terminal is assumed to have two on-line modems and one spare. The satellite weight (and cost) computed are overstated since only a trunking transponder is considered. The best optimization procedure would be multi-dimensional. The purpose here is to indicate a system approach which could support conclusions about technology development needs.

Table 6-11

Non-Regenerative Trunking Transponder  
(B.O. = 6/0)

$G/T$ (dB-°K)	$G/T_g$ (dB-°K)	D(20) (ft)	D(30) (ft)	Sat. Tx (W)	Term. Tx (W)	$w_c$ (lb)	$P_c$ (W)	$W_t$ (lb)	$P_t$ (W)
25.2	20.0	65.1	42.0	3.9	117	2685	2400	6830	3260
33.1	27.8	35.8	23.5	2.2	66	2030	1790	5340	2490
36.6	31.5	27.0	17.5	1.6	48	1815	1580	4845	2230
40.0	34.7	21.1	13.5	1.3	39	1680	1470	4550	2100
43.0	37.7	16.8	11.0	1.0	30	1575	1360	4300	1950
46.0	36.1	18.9	12.1	1.1	33	1627	1400	4410	2000

Table 6-12

## Trunking Terminal Satellite and System Cost

G/T (dB-°K)	D <sub>term</sub> (ft)	Terminal Cost (\$K)	Terminal Segment Cost (\$M, Q~80)	NRE Sat Cost	RE Sat Cost	Total System (\$M)
25.2	5.2	1360	76	221	114	411
33.1	13.5	1377	77	167	91	335
36.6	18.5	1402	79	149	83	311
40.0	30.0	1458	82	139	78	299
43.0	24.0	1555	87	130	74	291
46.0	33.0	1676	94	134	76	304

The lowest cost approach indicated for the trunking system is the use of a  $G/T = 43$  dB-°K terminal and a 17 ft/11 ft satellite antenna with 1 W power amplifiers. This conclusion for the trunking channels depends on the assumptions made. If the link margin assumption were to change from 3 dB (for dual diversity) to 13 dB (for dual diversity), the results would be likely to change. The minimum cost solution for the 3 dB margin requires a 43 dB-°K terminal. The additional 10 dB in margin is accomplished by an increase in satellite EIRP alone. This results in a change in  $D(20)$  to 36 feet,  $D(30)$  to 23 feet,  $p$  to 2.2 W, and  $P_{term}$  to 66 W. The transponder weight and power increases to 2,030 lb and 1,790 W, respectively. Satellite weight and power increases to 5340 lb and 2490 W; the non-recurring (NRE) investment is \$167M, the recurring (RE) satellite cost is \$91M, for a total system cost of \$345M. This is an increase of approximately \$53M over the 3 dB margin system. If the satellite is kept constant and the 10 dB improvement is obtained by increasing the terminal  $G/T_s$ , the results are quite different. This implies use of an 88 foot diameter antenna with the 50-°K low noise amplifier ( $G/T_s \approx 46$  dB-°K,  $G/T \approx 55$  dB-°K). The terminal (Q~1) now costs \$3.4M and the terminal segment (Q~80) now costs \$184M, an increase of \$98M over the 3 dB margin case. Thus, one concludes that an investment in a larger satellite is cost effective even when only 80 terminals are involved.

#### 6.6.1 Case (1) Trunking plus Non-regenerative TDMA Direct-to-User Transponder

In the comparison of various direct-to-user transponder types, the minimum cost trunking configuration (for the 3 dB margin case) is used. Table 6-13 presents the satellite implications for the case (1) direct-to-user transponder. Table 6-14 presents the terminal segment costs for the direct-to-user terminals. It is assumed that there is only one modem per terminal and that the total direct-to-user terminal population is on the order of 5,000.

Table 6-13

#### Non-Regenerative Direct-to-User TDMA Transponder (B.O. = 6/0)

G/T (dB-°K)	G/T <sub>s</sub> (dB-°K)	D(20) (ft)	D(30) (ft)	Sat. Tx (W)	Term. Tx (W)	W <sub>c</sub> (lb)	P <sub>c</sub> (W)
17.0	11.8	15.6	10.0	19.0	425	535	650
20.0	14.8	12.4	8.0	14.9	335	430	605
23.0	17.8	9.9	6.2	11.8	265	365	570
26.0	20.8	7.8	4.9	9.4	210	315	540
29.5	24.3	6.0	3.9	7.2	160	285	515
33.0	27.8	4.6	3.0	5.5	125	265	495
36.0	30.8	3.6	2.3	4.4	100	250	480

The cost of a satellite and satellite system as a function of direct-to-user terminal G/T is presented in table 6-15. This table does not include the trunking transponder or trunking terminals.

Table 6-14

Direct-to-User TDMA Terminal Costs  
Case (1)

G/T (dB-°K)	Ant. & LNA (\$K)	D (ft)	Tx & Up/Down Conv. (\$K)	Total RF (\$K)	Terminal Q=1 (\$K)	Total Q=5000 (\$M)
17.0	25	2.5	80	179	339	848
20.0	25	3.3	76	172	332	830
23.0	26	4.7	73	168	328	820
26.0	28	6.3	68	163	323	808
29.5	33	9.8	65	167	327	818
33.0	44	15.0	63	182	342	855
36.0	56	21.0	59	197	357	893

Table 6-15

Direct-to-User System Costs  
Case (1)

G/T (dB)	Satellite Costs		Terminal	Total
	NRE (\$M)	RE (\$M)	(\$M)	(\$M)
17.0	108	39	848	995
20.0	97	35	830	962
23.0	90	33	820	943
26.0	84	32	808	924
29.5	80	30	818	928
33.0	77	30	855	962
36.0	76	29	893	998

One can see that the minimum cost direct-to-user system implies the use of a G/T = 26 dB-°K terminal which is 6.3 feet in diameter, has a GaAs FET LNA, and a 210 W transmitter. The satellite would have D(20) = 7.8 feet, D(30) = 4.9 feet, and a 10 W TWT per beam. This holds only for the assumptions used. Consequently the minimum cost system for case (1) is as described in table 6-16.

Table 6-16

Minimum Cost System  
Case (1)

Item	Trunking Channel	Direct-to-User Channel
No. Beams (type)	40 (fixed)	2 (scanning)
Modulation	CQPSK (up/down)	CQPSK (up/down)
Access	TDMA	TDMA
Bandwidth/Beam	2400 MHz	100 MHz
Data Rate/Beam	3300 Mb/s	150 Mb/s
Sat. Antenna Size	16.8'/11.0'	7.8'/4.9'
Sat. Power/Beam	1 W	10 W
$W_c/P_c$	1575 lb/1370 W	315 lb/540 W
Terminal Size	24'	6.3'
Terminal Power	30 W	210 W
No. Terminals	80	5000
Terminal Cost	\$87M	\$808M
$W_c/P_t$	5215 lb/2635 W	
NRE	\$300M	
RE	\$ 89M	
Total Cost	\$1284M	

### 6.6.2 Case (2) Trunking plus Regenerative TDMA Direct-to-User Transponder

Table 6-17 illustrates the satellite implications for the case (2) direct-to-user transponder. Table 6-18 presents the costs for the case (2) direct-to-user terminals.

Table 6-17

Regenerative Direct-to-User TDMA Transponder  
(B.O. = 0/0)

G/T (dB-K)	G/T <sub>s</sub> (dB-K)	D(20) (ft)	D(30) (ft)	Sat Tx (W)	Term Tx (W)	$W_c$ (lb)	$P_c$ (W)
17.0	11.8	9.5	6.4	11.4	81	350	565
20.0	14.8	7.5	4.8	9.0	64	310	535
23.0	17.8	6.0	3.8	7.2	51	280	515
26.0	20.8	4.8	3.1	5.7	41	265	500
29.5	24.3	3.6	2.3	4.4	31	250	480
33.0	27.8	2.8	1.8	3.3	23	245	470
36.0	30.8	2.2	1.4	2.7	19	237	460

Table 6-18

Direct-to-User TDMA Terminal Costs  
Case (2)

G/T (dB-°K)	Ant. & LNA (\$K)	D (ft)	Tx & Up/Down Conv. (\$K)	Total RF (\$K)	Terminal Q=1 (\$K)	Total Q=5000 (\$M)
17.0	25	2.5	57	139	299	748
20.0	25	3.3	55	136	296	740
23.0	26	4.7	53	134	294	735
26.0	28	6.3	50	133	293	733
29.5	33	9.8	47	136	296	740
33.0	44	15.0	46	153	313	783
36.0	56	21.0	45	172	332	830

The cost of a satellite and satellite system (not including the trunking transponder or terminals) as a function of direct-to-user terminal G/T is presented in table 6-19 for case (2).

Table 6-19

Direct-to-User System Costs  
Case (2)

G/T (dB-°K)	Satellite Costs		Terminal (\$M)	Total (\$M)
	NRE (\$M)	RE (\$M)		
17.0	88	33	748	869
20.0	83	31	740	854
23.0	80	30	735	845
26.0	78	30	733	843
29.5	76	29	740	845
33.0	75	29	783	887
36.0	74	28	830	932

One can see for case (2) that the minimum cost direct-to-user system implies the use of a  $G/T = 26.0$  dB-K terminal as in case (1) except that the uplink transmitter is now 40 W. The satellite would have  $D(20) = 4.8$  feet,  $D(30) = 3.1$  feet, and a 6 W TWTA per beam. The minimum cost system for case (2) is described in table 6-20.

Table 6-20  
Minimum Cost System  
Case (2)

Item	Trunking Channel	Direct-to-User Channel
No. Beams	40 (fixed)	2 (scanning)
Modulation	CQPSK (up/down)	DQPSK/CQPSK
Access	TDMA	TDMA
Bandwidth/Beam	2400 MHz	100 MHz
Data Rate/Beam	3300 Mb/s	150 Mb/s
Sat. Antenna Size	16.8"/11.0'	4.8"/3.1'
Sat. Power/Beam	1 W	6 W
$W_c/P_c$	1575 lb/1360 W	265 lb/500 W
Terminal Size	24"	6.3"
Terminal Power	30 W	40 W
No. Terminals	40	5000
Terminal Cost	\$87M	\$733M
$W_c/P_c$	5105 lb/2585 W	
NRE	\$ 294M	
RE	\$ 87M	
Total Cost	\$1201M	

#### 6.6.3 Case (3) Trunking plus Regenerative FDMA Direct-to-User Transponder

Table 6-21 illustrates the satellite implications for case (3) and table 6-22 presents the terminal costs.



Table 6-21

**Regenerative Direct-to-User FDMA Transponder**  
(B.O. = 0/0)

G/T (dB-°K)	G/T <sub>s</sub> (dB-°K)	D(20) (ft)	D(30) (ft)	Sat. Tx (W)	Term Tx (W)	w <sub>c</sub> (lb)	P <sub>c</sub> (W)
17.0	11.8	9.5	6.4	11.4	8.1	350	565
20.0	14.8	7.5	4.8	9.0	6.4	310	535
23.0	17.8	6.0	3.8	7.2	5.1	280	515
26.0	20.8	4.8	3.1	5.7	4.1	265	500
29.5	24.3	3.6	2.3	4.4	3.1	250	480
33.0	27.8	2.8	1.8	3.3	2.3	245	470
36.0	30.8	2.2	1.4	2.7	1.9	237	460

Table 6-22

**Direct-to-User FDMA Terminal**  
Case (3)

G/T (dB-°K)	Ant. & LNA (\$K)	D (ft)	Tx Up/Down Conv. (\$K)	Total RF (\$K)	Terminal Q=1 (\$K)	Total Q=5000 (\$M)
17.0	25	2.5	37	105	205	513
20.0	25	3.3	35	102	202	505
23.0	26	4.7	35	104	204	510
26.0	28	6.3	33	104	204	510
29.5	33	9.8	32	111	211	528
33.0	44	15.0	31	127	227	568
36.0	56	21.0	30	147	247	618

The cost of a satellite and satellite system for case (3) (not including the trunking transponder or terminals) as a function of direct-to-user terminal G/T is presented in table 6-23.

The minimum cost direct-to-user system has a broad character to it. In case (3), the 20.0 dB-°K terminal is selected. It has a 3.3 foot antenna and a 6 W transmitter. The satellite would have D(20) = 7.5 feet, D(30) = 4.8 feet, and a 9 W TWT per beam. The minimum cost system for case (3) is described in table 6-24.

Table 6-23

**Direct-to-User System Costs**  
Case (3)

G/T (dB-°K)	Satellite Costs		Terminal (\$M)	Total (\$M)
	NRE (\$M)	RE (\$M)		
17.0	88	33	513	634
20.0	83	31	505	619
23.0	80	30	510	620
26.0	78	30	510	618
29.5	76	29	528	633
33.0	75	29	568	672
36.0	74	28	618	720

Table 6-24

**Minimum Cost System**  
Case (3)

Item	Trunking Channel	Direct-to-User Channel
No. Beams	40 (fixed)	2 (scanning)
Modulation	CQPSK (up/down)	DQPSK/CQPSK
Access	TDMA	FDMA
Bandwidth/Beam	2400 MHz	100 MHz
Data Rate/Beam	3300 Mb/s	150 Mb/s
Sat. Antenna Size	16.8'/11.0'	7.5'/4.8'
Sat. Power/Beam	1 W	9 W
$w_c/P_c$	1575 1b/1360 W	310 1b/535 W
Terminal Size	24'	3.3'
Terminal Power	30 W	6 W
No. Terminals	80	5000
Terminal Cost	\$87M	\$505M
$w_c/P_t$		5200 1b/2630 W
NRE		\$300M
RE		\$ 89M
Total Cost		\$981M

## 6.7 TECHNOLOGY IMPLICATIONS

An examination of the three cases yields the following conclusions. First, in the case of the trunking transponder, it is more cost effective to use a larger satellite antenna with a smaller output power amplifier. Thus, one would conclude that development money should go into larger antenna structure technology and into 20 GHz low to medium power (<5 W) space qualified GaAs FET amplifiers. This also leads to lower power transmitters for the ground segment terminals. It would be highly desirable to develop low cost medium power TWTAs (<50 W) at 30 GHz for the terminal segment, as well as low noise GaAs FET low noise amplifiers at 20 GHz. There is a need for the development of high rate TDMA modems (>800 Mb/s).

In the case of the direct-to-user channels, the following trend has been uncovered. On-board processing of higher rate TDMA signals yields a cost saving of approximately \$80M for the assumptions presented here. This is not as dramatic as the \$300M cost saving which results if FDMA uplinks are used in conjunction with on-board signal processing. The preliminary results indicate development of an FDMA based system concept with on-board signal regeneration for the direct-to-user terminals. This implies a need to develop scanning beam antennas and FDMA to TDM on-board signal processing techniques. Solid state (GaAs FET) terminal transmitters at modest power (5 W to 10 W) at 30 GHz also should be developed.

These preliminary conclusions are based on the assumptions and methodology presented. Further refinement will occur during the next phase of the study.

#### REFERENCES

BTL, "Deep Space Communications and Navigation Study," Vol. 2, Final Report, Bell Telephone Laboratories, May 1968.

Cuccia, C. L., "Status Report: Solid State Baseline of Telecommunications into the 1980's," MSN; April/May 1976.

Frediani, D. J., "Technology Assessment for Future MILSATCOM Systems; The 20/30 GHz Band," DCA-4, MIT Lincoln Laboratory; August 1978.

ITU, "Final Acts of the World Administrative Radio Conference for Space Telecommunications," International Telecommunications Union, Geneva; 1971.

Le Fande, R. A., "Attenuation of Microwave Radiation for Paths Through the Atmosphere," NRL Report No. 6766, November 1968.

"MILSATCOM Systems Architectures, Annex G 'Cost Models'," DCA/MSO Washington, DC, March 1976.

## SECTION 7

### CONCLUSIONS AND RECOMMENDATIONS

Conclusions and recommendations developed during this study are summarized in this section. The organization of topics follows the sequence in the preceding detailed discussions.

#### 7.1 MULTIPLE ACCESS/MULTIPLEXING (Section 3.2)

- Multiple beams are very attractive for providing frequency reuse and higher gain. This will greatly alleviate potential spectral crowding problems, will provide added link margin to combat rain attenuation, and will permit smaller, and, therefore, lower cost terminals.
- In the order of at least 100 coverage areas should be achieved with a combination of fixed and scanning beams. For example, approximately 20 fixed narrow beams on the large population centers and considerably fewer and wider scanning beams to handle remote areas would offer an interesting mixture capable of handling both trunking and direct-to-user traffic.
- If the frequency bands of a fixed beam and a scanning beam are disjoint, then the scanning beam can also be used in the heavily populated area to offload the fixed beam traffic on occasion.
- TDMA on the uplinks is appropriate for the large terminals operating in the fixed beams. These terminals would be capable of high burst rates and able to handle the necessary synchronization requirements. The satellite would operate in an SS-TDMA (RF switched) mode for these higher data rate (e.g., T3 and T4 carrier rates) trunking terminals.
- FDMA on the uplinks and TDM on the downlinks is preferred for the smaller terminals operating in the scanning beams. These terminals may not be capable of the burst rates and synchronization required to operate in a TDMA mode while in the dwell of a scanning uplink beam. Although a TDM mode in

a downlink scanning beam is feasible because the entire signal originates from a common platform, an FDM mode for the downlink scanning beams would also be possible. This would tend to equalize uplink and downlink burst rate disparities but could complicate spacecraft intermodulation problems. The satellite would operate in a regenerative (demodulation/remodulation) mode for these lower data rate (e.g., T1 and T2 carrier rates) direct-to-user terminals.

- An upper limit of 500 MHz or 600 MHz instantaneous bandwidth seems reasonable for a fixed beam based on projected digital clocking rates of about 300 MHz for readily available terminal hardware. The allocated 2.5 GHz bandwidth in the 30/20 GHz frequency spectrum can therefore be filled by four or five disjoint bandwidth sections as necessary. This would permit a modular satellite design which might simplify the system considerably.
- An upper limit of approximately 300 MHz instantaneous bandwidth appears likely for a scanning beam based on the 300 MHz clocking rate for the demodulators/remodulators and memory speeds required for processing data, reformatting, transmultiplexing, downlink routing, etc.
- Typical throughputs achievable with multiple beams for RF switched data are in the order of 10 Gb/s, perhaps more than adequate for anticipated trunking demands. Throughputs for regenerated data may be limited by on-board processing speeds in the order of 100 Mb/s.
- On-board digital switching requirements do not determine spacecraft technology. Brute force designs of even a 100 x 100 digital switch appear feasible with VLSI techniques. Non-blocking Clos networks of digital switches could be implemented to reduce complexity if desired.

## 7.2 MODULATION/DEMODULATION (Section 3.3)

### Conclusions

- The particular type of modulation employed by large trunking terminals operating in a TDMA or TDM mode is not a critical issue. QPSK or some variant such as offset QPSK or MSK is

preferable for bandwidth efficiency. Offset QPSK and MSK may have some advantage over QPSK in passing through system non-linearities.

- The type of modulation is more important for direct-to-user terminals relying on on-board regeneration. MSK or its continuous frequency variants are superior to QPSK or offset QPSK for reducing interchannel interference or allowing closer center frequency separations in an FDMA or FDM mode.

### Recommendations

- System Concepts

Further study of TDM switching on the ground as opposed to space division switching on-board the satellite is warranted to uncover possible advantages in spacecraft complexity and communications payload throughput efficiency.

Instead of using separate uplink and downlink frequency bands for non-interference purposes and frequency translation on-board the satellite, explore time translation (store and forward) for non-real time data using the same uplink and downlink frequency bands. This would alleviate frequency management problems for this kind of traffic by at least a factor of two and permit the same antenna structures to be used for both the uplinks and downlinks without any frequency interference.

Two geostationary orbit satellites with crosslinks, one over the eastern U.S. and one over the western U.S., should be considered and compared to the usual single geostationary satellite configuration. Increases in system availability and additional flexibility in on-board RF switching and memory management would be among the potential advantages.

The control structures of various hybrid networks composed from star nets, loops, and completely connected subnets should be investigated to determine cost tradeoffs as a function of the number of nodes in the context of satellite signal processing.

● Technology

Continue research on suitable linear amplifiers for satellite repeaters. This would permit FDMA/FDM with a non-regenerating transponder without transmitter power control at the terminals and the minimization of intermodulation problems on the downlink side of the satellite.

Further study of scanning beam implementation is strongly recommended to further establish and refine the system tradeoffs between fixed and scanning beams as a function of spacecraft weight and power.

Investigate SAW devices and CCDs for on-board demodulation and other signal processing at high data rates with SSMA to combat interchannel interference among simultaneous users. Support the development of spaceborne high capacity bubble memories for buffering and non-real time regeneration applications perhaps involving time translation. These developments could lead to lower cost communication packages on the satellite.

A more vigorous pursuit of high risk satellite technology necessary for very small inexpensive terminals is suggested to prepare for the far term direct-to-user market demand of millions of individual citizens.

● Analysis

Attempt a more comprehensive global optimization of the systems parameters including link power budgets and satellite and terminal cost factors.

Exploration of pseudonoise frequency hopping, linear FM (chirp), and possible hybrid combinations should continue, and include other more advanced modulation/demodulation techniques for performance improvements, cost reductions, and frequency spectrum conservation.



### 7.3 REGENERATION (Section 3.4)

- A satellite system with three different backoff states has been examined for uncoded and coded data. For all cases examined, the operating points for a given performance level are significantly lower for the regenerative repeater case than for the conventional repeater case. That is, the regenerative repeater results in either a greater fade margin or a reduced power requirement on both uplinks and downlinks. The improvement potential is greatest for the O/O HPA/TWTA backoff case, an operating state which is energy efficient but which suffers the largest degradation. This indicates that the regenerative repeater accomplishes its greatest improvement when the degradation encountered for the conventional case is the greatest.
- Performance was compared on the basis of minimum uplink and downlink ENRs; other operating points can be selected to achieve other objectives. For example, if an uplink ENR of 14 dB is selected, the resulting improvement in downlink ENR between HDP and nonregeneration is 8.5 dB. That is, the downlink ENR can be minimized at the expense of the uplink and still achieve the same performance.

### 7.4 BASEBAND PROCESSOR (Section 4)

- A digital processor capable of performing high speed switching of multiple T1 and T2 uplink/downlink channels appears to be feasible in the 1990 to 2000 time frame.
- Gigabit/sec digital device technology is under development. Work should be undertaken to develop a space qualified version.
- Multiprocessor system designs are being developed; however, the architectural configuration for efficient multiport memory with a large number of ports and digital switch interconnecting subsystems needs to be further developed.
- Digital programmable multiple high rate all-digital demodulator/modulator modules need to be developed.

- Designs with low power requirements and which can easily be checked for proper operation using built-in self test algorithms should be developed.
- Fault-tolerant design techniques should be applied in each area so as to maximize on-orbit lifetime and usefulness.

## 7.5 MICROWAVE SWITCH (Section 5)

### Conclusions

- This study shows that a large microwave switch can be built; some of the architectures and technology that can be used have been indicated and tentative specifications for an RF switch have been listed.
- Coupler cross bar architecture appears superior from the point of view of minimum weight and volume, broadcast capability, and reliability.
- Multiple beam phased arrays have too many technical problems to be considered for switching applications. They could provide another practical service, e.g., positioning a beam in some area which is not normally covered or augmenting coverage in areas which generate high communications traffic. The signals received or transmitted would be routed by the switch.
- Ferrite devices are too slow, too heavy, and require too much power to be useful in a large switch.
- Optical switches are not presently smaller than microwave switches and will remain so because of limitations on the ability to make large controlled changes to the index of refraction of optical materials. They are fast and require biasing of up to several hundred volts for operation. Signals must be transduced from microwave to optics and back to microwave, which increases component expense.
- PIN diodes are a mature technique and can do the job. There have been improvements in drive power consumption but drivers will be required. PINs are no longer the fastest

elements. They require more microwave and biasing circuitry than do FETs.

- FETS can exceed the performance of PINs as an RF switch device and could be the best devices if optimized.
- Switch design to provide the necessary microwave signal, control, mounting, and thermal cooling paths, and to provide for assembly and maintenance before (perhaps after) launch needs further research. A common concern is switch assembly. Normal three dimensional wiring techniques used in switching at lower frequencies for signals are no longer applicable at microwave.

#### Recommendations

- Design Study

In view of the problem of switch layout at microwave, a switch design study should be carried out. It should determine layout, packing, and construction practice for a microwave cross bar (function) switch of (100 x 100) size. The study should cover routing of microwave and control signals, methods of cooling, interconnection, fabrication, and assembly. The study should assume use of devices, architecture, and transmission media recommended in this report. The study should cover logic control, reliability and one-to-many broadcasting.

- Media

Use of coupler cross bar architecture with feed throughs realized in microstrip should be considered for the switch. This configuration affords the smallest size and weight and provides the highest reliability.

- Devices

There are two usable devices, PINs and FETs. FETs are so promising that, despite needed optimization for switching, they are the preferred device. A development program should be instituted with recognized FET suppliers to develop optimum FETs for switching.

Because of the need to transduce into and out of the optical medium, optical switching should be reviewed in about five years to see if a breakthrough makes it competitive.

Phased arrays and ferrites are unsuitable.

- Switch Size

A 100 x 100 microwave cross bar switch is a necessary and useful element in a TDMA satellite communications system in view of the 73 municipalities in the U.S. with more than a half million population plus the possibility of using several beams for the largest municipality.

## 7.6 COST ANALYSIS (Section 6)

- Trunking Transponder

Use of a larger satellite antenna with a smaller output power amplifier is more cost effective. Development money should go into larger antenna structure technology and 20 GHz low to medium power (<5 W) space qualified GaAs FET amplifiers. This leads to lower power transmitters for the ground segment terminals. Low cost medium power TWAs (<50 W) at 30 GHz for the terminal segment, as well as low noise GaAs FET low noise amplifier at 20 GHz, should be developed. There is a need to develop high rate TDMA modems (>800 Mb/s).

- Direct-to-User Channels

On-board processing of higher rate TDMA signals yields a cost saving of approximately \$80M for the assumptions considered. This is nowhere near as dramatic as the \$300M cost saving which results if FDMA uplinks are used in conjunction with on-board signal processing. Development of an FDMA based system concept with on-board signal regeneration for the direct-to-user terminals is indicated. This implies a need for development of scanning beam antennas and FDMA to TDM on-board signal processing techniques. Solid state (GaAs FET) terminal transmitters at moderate power (5 W - 10 W) at 30 GHz also should be developed.

## APPENDIX A

### DIGITAL SWITCHING NETWORKS

Configurations and possible implementations of  $m \times n$  digital switches are discussed in this appendix. A straightforward realization of an  $m \times n$  rectangular switch with OR gates and AND gates is explained first. Simpler realizations, in the sense of requiring fewer circuit elements and control lines, composed of cascaded interconnected stages of smaller sized switches are then presented. Although these multistage realizations may introduce some degree of blocking, a low blocking probability may be worth the savings in switch complexity.

Figure A-1 represents a canonical  $m \times n$  switch requiring  $n$   $m$ -input OR gates and  $mn$  2-input AND gates. If necessary each  $m$ -input OR gate could be replaced by  $R^{L-1}$  to  $R^{L-1}$   $R$ -input OR gates, where  $L$  is the least integer no less than  $\log_2 m$ . Complexity varies as  $mn$  if it is measured as being proportional to the number of AND gates (crosspoints), OR gate inputs, and/or control lines. This measure has implications regarding the necessary sizes and speeds of switches and control memory, etc., implemented from LSI chips.

The switch operates as follows. Suppose the bit sequence signal on input line  $i$  is to be transferred to output line  $j$ . Then control line  $ij$  is made logical 1 to activate  $AND_{ij}$ . The  $n-1$  other control lines designated  $ik$ ,  $k \neq j$  are at logical 0. Furthermore, the  $m-1$  other control lines designated  $kj$ ,  $k \neq i$  are at logical 0. Thus, each input line is connected to one and only one distinct output line, and each of the  $\min(m, n)$  activated AND gates passes the binary data signal essentially without change from the associated input to the designated output. There is no blocking. In other words, for an arbitrary pair of free input and output lines which are to be connected, there is always a way to connect them without disturbing or rearranging any other previously made connections.

Let the switching network complexity be defined as the total number of control lines. Thus, a single  $m \times n$  switch has complexity  $mn$ . The remainder of this appendix is divided into two parts which discuss canonical switching networks with the same number of inputs and outputs but with complexity smaller than  $m^2$ . The first part discusses multistage networks of square switches based on a factorization of  $m$ . For this class of network, it is shown that the

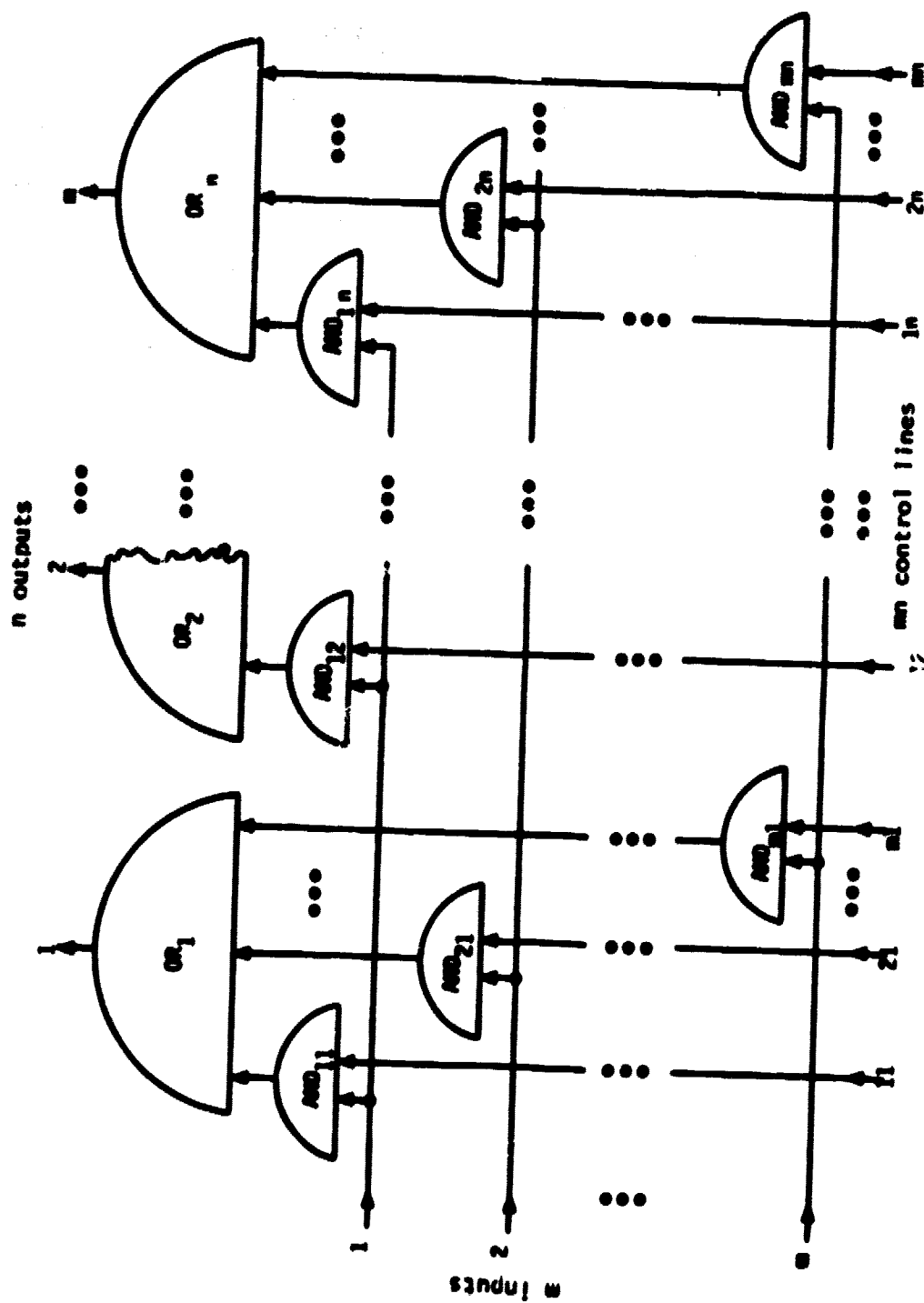


Figure A-1. Canonical  $m \times n$  Digital Switch Realization

minimum complexity results from choosing each factor to be the  $s$ th root of  $m$  for an  $s$ -stage network. Unfortunately, these networks are blocking. The second part discusses certain non-blocking multistage networks that may contain rectangular switches. Although the complexity will be increased over that of the first class, it will still be less than  $m^2$ , the complexity of a single  $m \times m$  switch.

#### MINIMAL COMPLEXITY BLOCKING NETWORKS OF SQUARE SWITCHES

For simplicity the following multistage realizations of an  $m \times m$  switch are restricted to compositions of square subswitches, where the outputs of each subswitch can feed another subswitch but only on the next stage. Such switching networks will have some positive blocking probability and will not necessarily be of an absolute minimum complexity for a given number of stages; equal propagation delays will be retained and some nice symmetries can result.

A canonical 2-stage realization of  $m \times m$  switch is shown in figure A-2. Here  $m$  is factored into the product  $m = m_1 m_2$ . There are  $m_2$   $m_1 \times m_1$  switches and  $m_1$   $m_2 \times m_2$  switches. It is easy to show that the optimum 2-stage network is satisfied by  $m_1 = m_2 = m^{1/2}$  in the sense that the overall switch complexity is minimized:

$$0 = \frac{d}{dm_1} (m_2 m_1^2 + m_1 m_2^2) = \frac{d}{dm_1} (m m_1 + m^2 / m_1) \\ = m - m^2 / m_1^2 \quad \text{which implies} \quad m_1 = m^{1/2}$$

and a complexity of  $2m^{3/2}$ , an improvement of a factor of  $(1/2)m^{1/2}$  over the switch of figure A-1. This suggests that in decomposing a given size switch into a 2-stage network, one should take  $m_1$  to be the integer as close to  $m^{1/2}$  as possible if  $m$  is not a perfect square.

However, the 2-stage network of figure A-2 has a high blocking probability. For example, at most one connection can be made between the  $j$ th input switch and the  $i$ th output switch. When such a connection is made, no other input of the  $j$ th input switch can be connected to an output of the  $i$ th output switch.

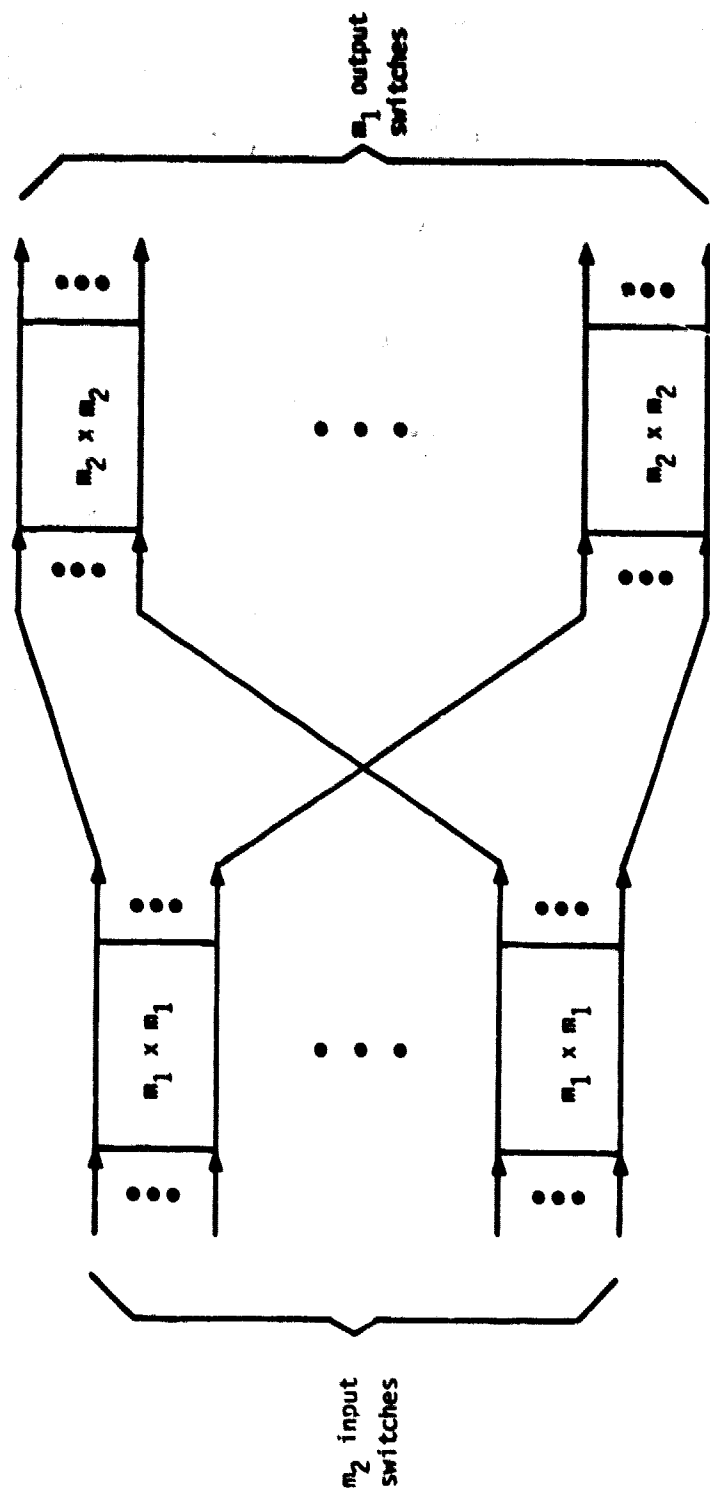


Figure A-2. Canonical 2-Stage Realization of an  $m \times m$  Digital Switch



One step further, a canonical 3-stage realization is shown in figure A-3. Now  $m = m_1 m_2 m_3$ , and the minimum complexity form occurs when  $m_1 = m_2 = m_3 = m^{1/3}$ . This result is obtained by solving the following pair of equations in two (arbitrarily selected) independent variables:

$$\begin{aligned}
 0 &= \frac{\partial}{\partial m_1} (m_2 m_3 m_1^2 + m_1 m_3 m_2^2 + m_1 m_2 m_3^2) \\
 &= \frac{\partial}{\partial m_1} (m m_1 + m m_2 + m^2 / m_1 m_2) \\
 &= m - m^2 / m_2 m_1^2 \text{ which implies } m_1^2 = m / m_2 = m_1 m_3 \text{ or } m_1 = m_3 \\
 0 &= \frac{\partial}{\partial m_2} (m m_1 + m m_2 + m^2 / m_1 m_2) \\
 &= m - m^2 / m_1 m_2^2 \text{ which implies } m_2^2 = m / m_1 = m_2 m_3 \text{ or } m_2 = m_3
 \end{aligned}$$

which imply that  $m_1 = m_2 = m_3 = m^{1/3}$  and a complexity of  $3m^{4/3}$ , a further improvement by a factor of  $(2/3)m^{1/6}$ . Again, in applying this result, one should choose the three factors of  $m$  to be as nearly equal as possible in the integers.

Extending this development in a straightforward manner to an  $s$ -stage network (see below), one can obtain a rule for interswitch connection analogous to the notes of figures A-2 and A-3, and a minimum complexity of  $sm^{(s+1)/s}$  with  $s$  columns of  $m^{(s-1)/s}$   $m^{1/s} \times m^{1/s}$  subswitches in each column (or approximately so with solutions in the integers). Note that the complexity is lower bounded by  $sm$  for large  $s$ , since  $\lim_{s \rightarrow \infty} m^{(s+1)/s} = m$ . The exact complexity  $c$  for

$$m = \prod_{i=1}^s m_i \quad (\text{A.1})$$

is given by

$$c = m \sum_{i=1}^s m_i. \quad (\text{A.2})$$

Specific cases of interest are given in table A-1.

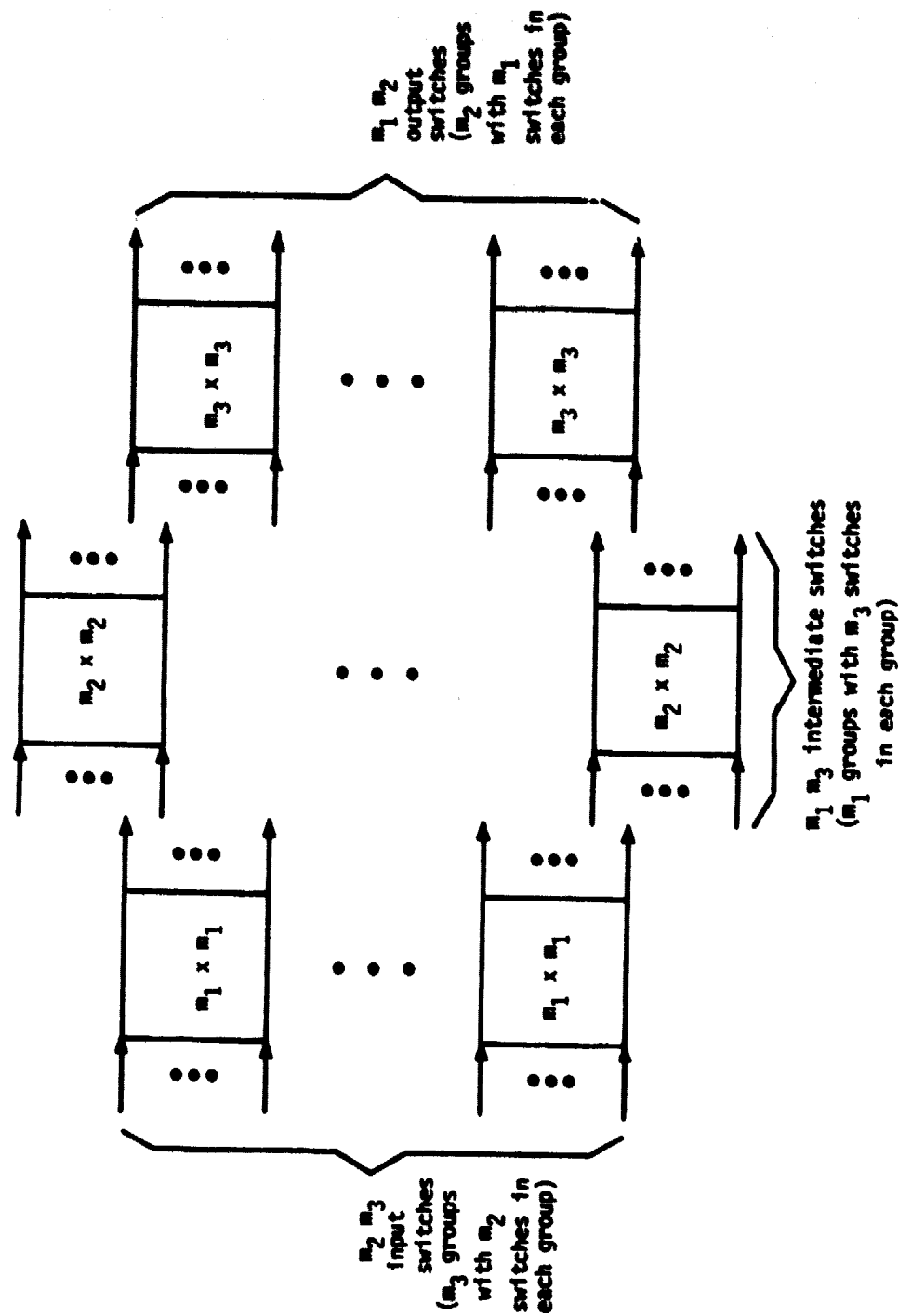


Figure A-3. Canonical 3-Stage Realization of an  $m \times m$  Digital Switch

Table A-1

Complexity (Number of Control Lines) in Various Implementations of an  $m \times m$  Digital Switch as an  $s$ -Stage Network of Individual Sub-switches of the Form of Figure A-1.

$m \backslash s$	1	2	3	4
36	1,296	432 ( $m_1=m_2=6$ )	360 ( $m_1=m_3=3$ $m_2=4$ )	360 ( $m_1=m_4=2$ $m_2=m_3=3$ )
72	5,184	1,224 ( $m_1=8, m_2=9$ )	1,008 ( $m_1=m_3=3$ $m_2=8$ )	864 ( $m_1=m_4=3$ $m_2=4, m_3=2$ )
100	10,000	2,000 ( $m_1=m_2=10$ )	1,400 ( $m_1=m_3=5$ $m_2=4$ )	1400 ( $m_1=m_4=2$ $m_2=m_3=5$ )
360	129,600	13,680 ( $m_1=18, m_2=20$ )	7,920 ( $m_1=5, m_2=9, m_3=8$ ) ( $m_1=m_3=6$ $m_2=10$ )	6,840 ( $m_1=m_4=6$ $m_2=2, m_3=5$ ) ( $m_1=m_4=3$ $m_2=5, m_3=8$ )
1,024	1,048,576	65,536 ( $m_1=m_2=32$ )	32,768 ( $m_1=m_3=8$ $m_2=16$ )	24,576 ( $m_1=m_4=4$ $m_2=m_3=8$ )

As previously mentioned, the multistage realizations discussed here are blocking in the sense that the existence of a free end-to-end input/output pair does not guarantee a path connecting this pair that is disjoint from every other path connecting occupied end-to-end input/output pairs. The computation of blocking probabilities for these switching networks is beyond the scope of this report. However, it is conjectured that significant reductions in complexity can justify low blocking probabilities in many cases.

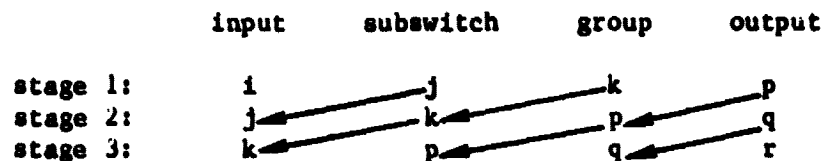
A proof that the minimum complexity is achieved with equal size subswitches for the canonical  $s$ -stage network case follows. With eqs. (A.1) and (A.2) and the LaGrange multiplier technique, the minimum complexity results from

$$0 = \frac{\partial}{\partial m_i} (\lambda (m - \sum_j m_j) + m \sum_j m_j) \\ = -\lambda + \sum_{j=1}^s m_j + m$$

which implies that

$$m_i = \lambda = m^{1/s}, \quad i = 1, \dots, s.$$

A general but not necessarily unique rule for specifying interswitch connections is now outlined along with a sketch of a proof showing the existence of a path connecting any input and any output ignoring problems of blocking. First, consider the case  $s=3$  and the following array showing the indices of the inputs, subswitches, groups, and outputs associated with an arbitrary chain of subswitches of an  $m \times m$  decomposition:



Here the  $p$ th output of the subswitch in stage 1 feeds the  $j$ th input of the subswitch in stage 2, and the  $q$ th output of the subswitch in stage 2 feeds the  $k$ th input of the subswitch in stage 3. Note that the connection between input and output in any subswitch is completely free to be selected, as in any 1:1 mapping.

Now suppose one desires to connect input  $i$  of subswitch  $j$  in group  $k$  of stage 1 to output  $w$  of subswitch  $u$  of group  $v$  in stage 3, where  $i, j, k$  and  $u, v, w$  are completely arbitrary and independent. From the above array it can be verified that one can achieve this by choosing  $p=u$ ,  $q=v$ , and  $r=w$ , since  $p$ ,  $q$ , and  $r$  are completely free variables. This proof can be generalized to any number of stages  $s$  by extending the array in both directions, creating super groups of subswitches in each stage, and observing the left shifts of the indices as one progresses to the next stage, as indicated by the arrows.

It is observed that by relaxing the implicit constraints of square subswitches and connecting a given subswitch output only to the input of a subswitch in the next column, simpler  $s$ -stage cascades may result. This is a fascinating subject in itself [Opferman, Tsao-Wu, 1971]. For a discussion of blocking but rearrangeable networks composed only of  $2 \times 2$  subswitches, where a given subswitch does not necessarily feed another subswitch on the next stage, see section 5 of this report. The next subsection of this appendix considers non-blocking canonical forms where the subswitches may be rectangular.

#### CANONICAL NON-BLOCKING REALIZATIONS WITH RECTANGULAR SWITCHES

In this subsection, canonical 3- and 5-stage networks of the Clos type are considered [Clos, 1953; Benes, 1965]. An odd number of stages is favored, evidently because non-blocking configurations are easier to achieve than with an even number of stages. This is also suggested by closer examination of specific examples of the 2- and 3-stage networks of the previous subsection.

A non-blocking 3-stage  $m \times m$  network is shown in figure A-4. The integer  $m_1$  is a non-trivial (not equal to 1 or  $m$ ) but otherwise arbitrary factor of  $m$ . The first stage contains  $m/m_1$   $m_1 \times (2m_1-1)$  rectangular subswitches, and the last stage contains the same number of  $(2m_1-1) \times m_1$  rectangular subswitches. The middle stage contains  $2m_1-1$  square subswitches of order  $m/m_1$ .

Consider the following worst case situation to verify the non-blocking property. Without loss of generality, suppose all but one of the inputs of subswitch 1 in stage 1 are connected to network outputs via any set of  $m_1-1$  subswitches in stage 2. Furthermore, suppose that all but one of the outputs of subswitch 1 in stage 3 are connected to network inputs via a set of  $m_1-1$  switches disjoint from the first set. Even in this case, there still exists a subswitch in stage 2 that is not in either of the first two sets

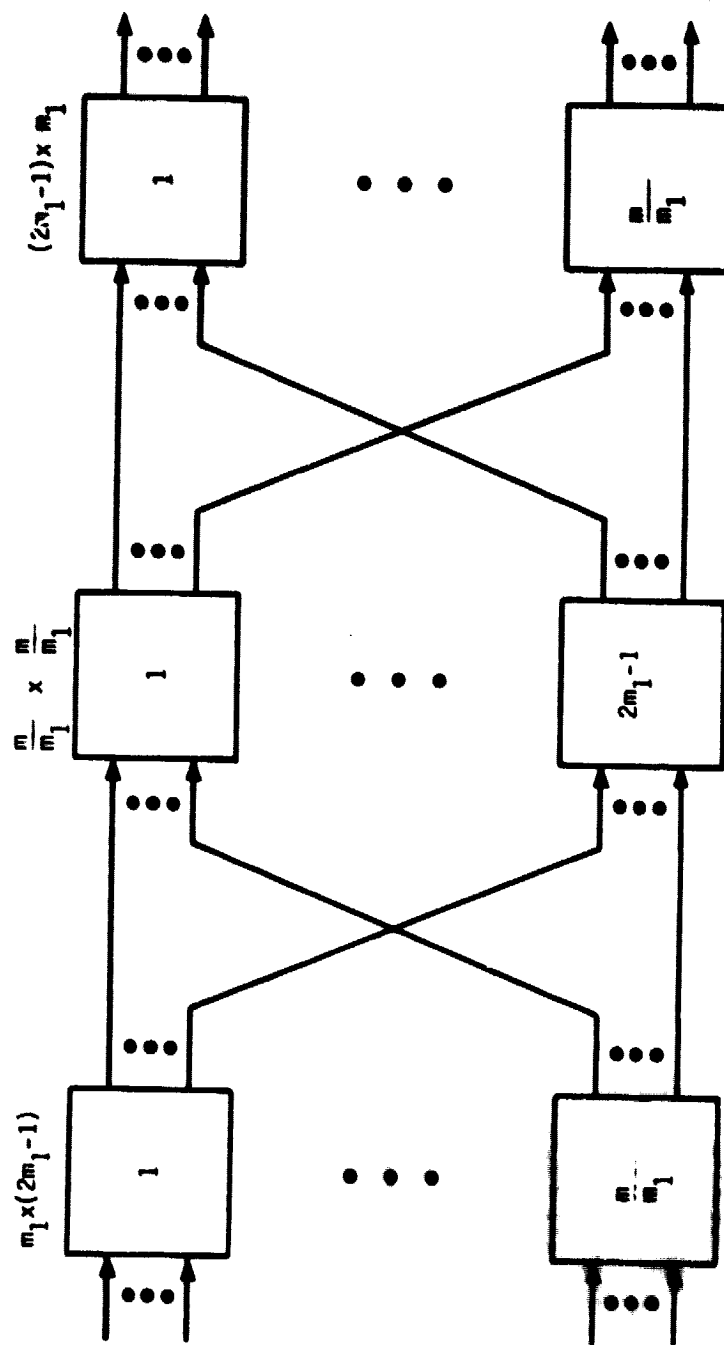


Figure A-4. Non-Blocking 3-Stage Clos Realization of an  $m \times m$  Digital Switch

which permits a connection between the remaining free input of subswitch 1 of stage 1 and the remaining free output of subswitch 1 of stage 3. This follows because there are  $2(m_1-1) + 1 = 2m_1-1$  subswitches in stage 2; with fewer subswitches, blocking would be possible.

This non-blocking construction can be extended to any  $s$ -stage network where  $s$  is an odd integer larger than three. For example, a 5-stage network is depicted in figure A-5. Here the middle stage contains  $2m_1-1$  levels each containing a non-blocking 3-stage network of the form just discussed. The 5-stage network is non-blocking because there are sufficient middle levels for the worst case input/output connections, as explained previously for the 3-stage network, and because each level is non-blocking in itself, by construction.

The complexities of these Clos networks is now compared for the same values of  $m$  used in the previous subsection. The best 3-stage Clos network has a smaller complexity than the best 5-stage Clos network for the smaller values of  $m$ . However, since these networks are non-blocking, their complexities will exceed those computed for the corresponding blocking realizations of table A-1.

The complexity of the 3-stage network of figure A-4 is

$$\begin{aligned} c_3(m) &= 2(m/m_1)m_1(2m_1 - 1) + (2m_1 - 1)(m/m_1)^2 \\ &= (2m_1 - 1)(2m_1^2 + m)m/m_1^2. \end{aligned} \quad (A.3)$$

The complexity of the 5-stage network of figure A-5 is

$$\begin{aligned} c_5(m) &= 2(m/m_1)m_1(2m_1 - 1) + (2m_1 - 1)c_3(m/m_1) \\ &= (2m + c_3(m/m_1))(2m_1 - 1) \\ &= (2m + ((2m_2 - 1)(2m_2^2 + m/m_1)(m/m_1))/m_2^2)(2m_1 - 1) \\ &= (2m_1 - 1)(2m_1m_2^2(m_1 + 2m_2 - 1) + (2m_2 - 1)m/(m_1m_2)^2) \end{aligned} \quad (A.4)$$

where  $m_2$  is an arbitrary non-trivial factor of  $m/m_1$ . Examples are given in tables A-2 and A-3. Note that the 5-stage networks have smaller complexity for  $m = 360$  and  $1024$ .

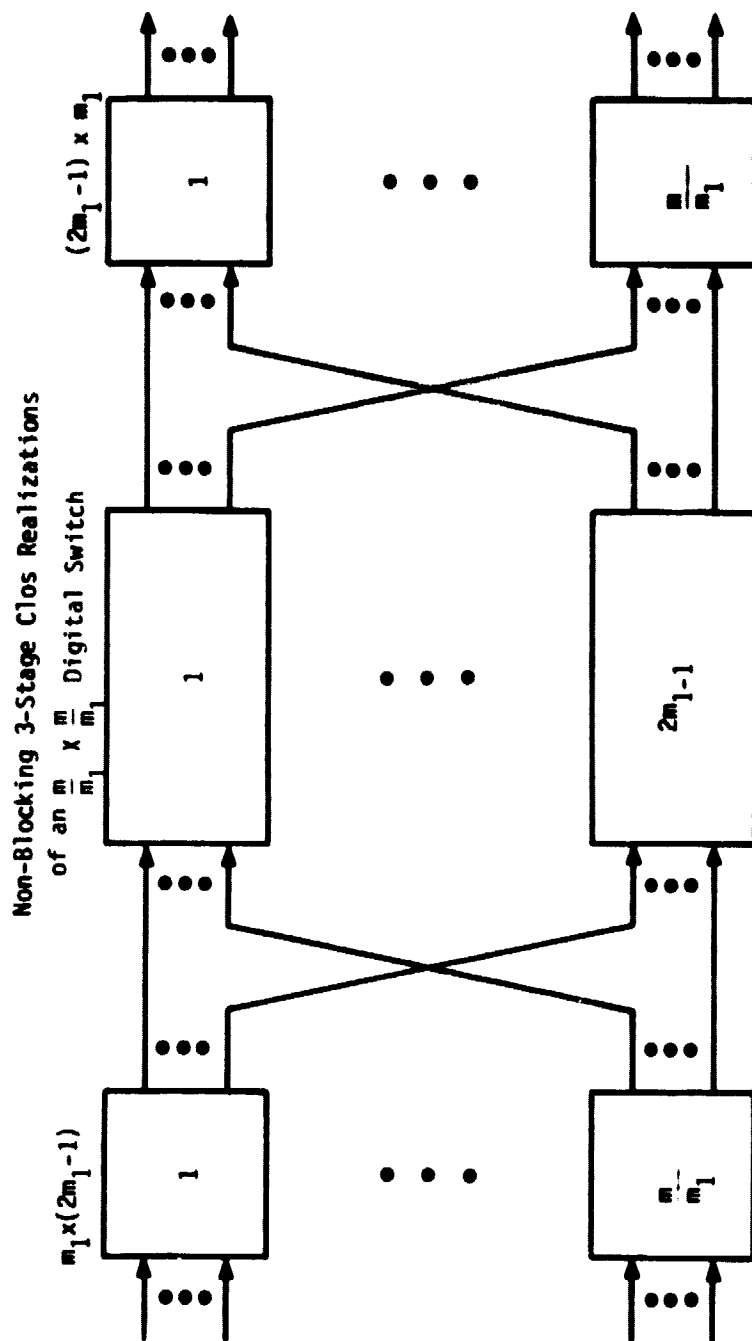


Figure A-5. Non-Blocking 5-Stage Clos Realization of an  $m \times m$  Digital Switch



Table A-2

Complexities of Various Non-Blocking 3-Stage  
Clos Realizations of an  $m \times m$  Digital Switch

$m$	$m_1$	$c_3(m)$
36	6	1,188
72	8	3,375
	9	3,536
100	10	5,700
360	18	39,200
	20	40,716
1,024	32	193,536

Table A-3

Complexities for Various Non-Blocking 5-Stage  
Clos Realizations of an  $m \times m$  Digital Switch

$m$	$m_1$	$m_2$	$c_5(m)$
36	3	3	1,360
72	4	3	3,528
	3	4	3,660
100	5	4	5,895
	4	5	6,125
360	8	5	33,885
	9	5	34,272
	5	8	36,855
1,024	8	8	145,920

It is known that

$$\min c(km) \leq (2k-1)(2km + \min c(m)) \quad (A.5)$$

where the minimisations are over all possible non-blocking Clos networks of the indicated orders [Moore, 1971]. Applying inequality (A.5) to tables A-2 and A-3, taking the best listed value of  $c(m)$  for  $\min c(m)$ , one cannot conclude from these data that there exists a Clos network of complexity smaller than 33,885 for order 360. In other words, the right-hand side of inequality (A.5) with  $c_3(m)$  for  $\min c(m)$  exceeds the value for  $c_3(360)$  in table A-3 for the cases  $(k,m) = (10,36)$  and  $(5,72)$ .

In the future, architecture involving TDM switching networks should also be given more consideration [Field, 1977].

#### REFERENCES

Clos, C., "A Study of Non-Blocking Switching Networks," Bell System Tech. J., Vol. 32, March 1953, pp. 406-424.

Beneš, V. E., Mathematical Theory of Connecting Networks and Telephone Traffic, Academic Press, New York, 1965.

Field, R. G., "Construction of Multistage Time-Division Switching Networks," Nat. Telecommun. Conf., Los Angeles, 5-7 December 1977, pp. 46.4-1 through 46.4-4.

Moore, E. V., private communication, 1971.

Opferman, D. C., Tsao-Wu, M. T., "On a Class of Rearrangeable Switching Networks Part I: Control Algorithm," Bell System Tech. J., May-June 1971, pp. 1579-1600.

## APPENDIX B

### MULTIPLE BEAM OPTIMIZATION

Assume a satellite solution involving both fixed and scanning beams, where the fixed beams handle RF-switched T3 and T4 trunking traffic while the scanning beams handle digitally-switched T1 and T2 direct-to-user traffic. This appendix examines the relationships among the bandwidths, numbers of beams, channels, bits per burst, etc., where throughput is to be maximized and satellite complexity such as RF switch size and bit processing speed are to be minimized. This provides a rationale for several interesting multiple beam examples until a more general optimization involving satellite and terminal costs can be developed.

Let the subscripts 1, 2, 3, 4 refer to four groups of users, nominally operating at the data rates of:

$$R_1 = 1.544 \text{ Mb/s}$$

$$R_2 = 6.312 \text{ Mb/s}$$

$$R_3 = 44.736 \text{ Mb/s}$$

$$R_4 = 276.176 \text{ Mb/s}$$

corresponding to T1, T2, T3, T4 carriers. The burst to data rate or frame to burst duration ratio is defined as

$$r_1 = B_1/k_1 = F_1/T_1. \quad (B.1)$$

The first two groups operate in scanning beams with FDMA, while the second two groups operate in fixed beams and basically employ TDMA, although more than one TDMA carrier may be allowed in a given fixed beam. The number of simultaneous beams is

$$M = \sum_{i=1}^4 M_i \quad (B.2)$$

and the number of coverage areas is given by

$$A = M_1 r_1 + M_2 r_2 + M_3 + M_4. \quad (B.3)$$

If  $N_i$  is the number of FDMA channels in group  $i$ , then the number of simultaneous users is

$$m = \sum_{i=1}^4 m_i = \sum_{i=1}^4 M_i N_i. \quad (B.4)$$

The number of concurrent users in one frame interval  $F_i$  is  $m_i r_i$ , where  $r_i$  is the number of disjoint areas covered by a scanning beam or the number of slots per frame in a fixed beam, and equal dwell times ( $T_1$  or  $T_2$ ) in each scanning beam and equal time slots ( $T_3$  or  $T_4$ ) in each fixed beam are assumed. Thus, the total number of concurrent users is

$$n = \sum_{i=1}^4 n_i = \sum_{i=1}^4 m_i r_i. \quad (B.5)$$

Since the throughput for group  $i$  is  $n_i R_i$ , the total satellite throughput is

$$t = \sum_{i=1}^4 t_i = \sum_{i=1}^4 n_i R_i. \quad (B.6)$$

If  $b_i = B_i T_i$  is the number of bits per burst for a group  $i$  user, then the total number of bits of data processed by the satellite on a burst basis is

$$q = \sum_{i=1}^4 q_i = \sum_{i=1}^4 b_i m_i = \sum_{i=1}^4 b_i M_i N_i \quad (B.7)$$

using eq. (B.4). (Note that the required sizes for the digital and RF switches are  $m_1+m_2$  and  $m_3$  and  $m_4$ , respectively, if crosslinking is performed only for T1 and T2 users.)

If  $\beta_1$  is the bps per Hz packing density assumed for FDMA in group 1, then the instantaneous bandwidth occupied by group 1 is

$$W_1 = B_1 N_1 / \beta_1. \quad (B.8)$$

It is assumed that simultaneous beam footprints are sufficiently separated that complete frequency reuse is attained in group 1. Furthermore, for disjoint frequency subbands for each beam class, the total instantaneous bandwidth is

$$W = \sum_{i=1}^4 W_i. \quad (B.9)$$

The system designer may be interested in maximizing throughput under the constraints on numbers of simultaneous beams, number of coverage areas, and instantaneous bandwidths, i.e., maximizing eq. (B.6) subject to eqs. (B.2), (B.3), and (B.9). Using the technique of Lagrange multipliers, for example, one can obtain conditions on the numbers of beams and the bandwidths by differentiating the objective function with respect to  $W_j$  and  $M_j$ , respectively. Since  $t_1 = n_1 R_1 = m_1 r_1 R_1 = B_1 N_1 M_1 = \beta_1 M_1 W_1$ , by eqs. (B.1), (B.4), (B.5), (B.6), and (B.8), one can write

$$\begin{aligned} 0 &= \frac{d}{dW_j} (\lambda_W (W - \sum_1 W_1) + \sum_1 \beta_1 M_1 W_1) \\ &= -\lambda_W + \beta_j M_j, \quad j = 1, 2, 3, 4 \end{aligned}$$

which along the eq. (B.2), implies

$$M_j = M / \beta_j \sum_{i=1}^4 \beta_i^{-1}, \quad j = 1, 2, 3, 4. \quad (B.10)$$

This means that for a given set of  $\beta_j$ 's, the number of simultaneous beams in group  $j$  should be inversely proportional to the bandwidth efficiency factor  $\beta_j$  of the modulation for that group. But this suggests that throughput will not be maximized. Indeed, throughput could be maximized by devoting all the beams and the entire bandwidth to a group with the largest bandwidth efficiency. Unfortunately, that solution is less interesting since both scanning and fixed beams are being investigated. Thus, one may be forced to accept considerably less than the true maximum throughput  $t_{\max} = \max_j (\beta_j) MW$ . Equation (B.10) therefore represents a critical point, but a saddle point and not a minimum, since zero throughput could be approached by pairing small bandwidths with large numbers of beams and large bandwidths with small numbers of beams, while still satisfying the constraints on  $M$  and  $W$ . Nevertheless, a saddle point may provide a reasonable solution in that a respectable throughput is obtained with an appropriate mix of services to all four groups.

Similarly, one can write

$$0 = \frac{d}{dM_j} (\lambda_M (M - \sum_1 M_1) + \lambda_A (A - M_1 r_1 - M_2 r_2 - M_3 - M_4) + \sum_1 \beta_1 M_1 W_1)$$

$$= \begin{cases} -\lambda_M - r_1 \lambda_A + \beta_j W_j, & j = 1, 2 \\ -\lambda_M - \lambda_A + \beta_j W_j, & j = 3, 4 \end{cases}$$

which implies a critical point for

$$W_4/W_3 = \beta_3/\beta_4$$

and

$$W_2/W_1 = \beta_1 (\lambda_M + r_2 \lambda_A) / \beta_2 (\lambda_M + r_1 \lambda_A).$$

For simplicity, it is assumed that the burst to data rate ratio is identical for groups 1 and 2, i.e.,

$$r_2 = r_1. \quad (B.11)$$

Again, it can be verified that a saddle point results:

$$W_2/W_1 = \beta_1/\beta_2 \text{ and } W_4/W_3 = \beta_3/\beta_4. \quad (\text{B.12})$$

Note that eqs. (B.9) and (B.12) do not specify how bandwidth should be divided between the T1-T2 groups and the T3-T4 groups.

The system designer may also be interested in minimizing the total burst rate of the T1 and T2 signals seen by the satellite demodulators. One would like to minimize

$$\begin{aligned} \sum_{i=1}^2 B_i m_i &= \sum_{i=1}^2 B_i M_i N_i = \sum_{i=1}^2 \beta_i M_i W_i \\ &= M(W_1 + W_2) / \sum_1 \beta_i^{-1} \\ &= M(\beta_1 + \beta_2) W_1 / \beta_2 \sum_1 \beta_i^{-1} \end{aligned} \quad (\text{B.13})$$

where eqs. (B.4), (B.8), (B.10) and (B.12) were applied. Since the total number of simultaneous beams  $M$  and the  $\beta_i$ 's are given, this implies that the T1 and T2 bandwidths should be minimized.

However, since  $B_i m_i = m_i r_i R_i = n_i R_i = t_i$ , from eqs. (B.1), (B.5) and (B.6), this means no on-board demodulation and zero throughput for the T1 and T2 groups. This trivial result is of little interest. Therefore, eq. (B.13) suggests that  $t_1 + t_2$  be held to an acceptable total burst rate. (Alternatively, one desires solutions with fair numbers of T1 and T2 users.)

Suppose the on-board memory speed is 10 nsec and bits are processed as 8-bit bytes moved serially in and out of memory. Then  $4 \times 10^8$  bps can be handled. For example, if  $M = 24$  and if  $\beta_1 = \beta_2 = 1/2$  and  $\beta_3 = \beta_4 = 1$ , then from eqs. (B.12) and (B.13)

$$W_1 = W_2 = 4 \times 10^8 \text{ bps } (1/2)(6)/24(1/2 + 1/2) = 50 \text{ MHz}$$

i.e., as much as 100 MHz could be comfortably devoted to the scanning beams. Note from eq. (B.10) that there would be eight scanning beams devoted to group 1 and eight to group 2 in this instance.



Alternatively, suppose one attempts to minimize the number of control lines  $(m_1 + m_2)^2$  for a digital switch of size  $m_1 + m_2$  for groups 1 and 2. (See appendix A.) From eqs. (B.2) and (B.4)

$$0 = \frac{d}{dM_1} \left( \sum_{i=1}^2 M_i N_i \right)^2 = 2(M_1 N_1 + M_2 N_2)(N_1 - N_2)$$

and similarly,

$$0 = \frac{d}{dM_2} \left( \sum_{i=1}^2 M_i N_i \right)^2 = (M_1 N_1 + M_2 N_2)(N_2 - N_1)$$

which imply that

$$N_1 = N_2 \quad (B.14)$$

for a non-trivial solution. Similarly, differentiation with respect to  $N_1$  and  $N_2$  suggests that

$$M_1 = M_2 \quad (B.15)$$

by symmetry. Substituting eqs. (B.14) and (B.15) into the objective function, one obtains  $4M_1^2 N_1^2$ . Unfortunately, the original objective function can assume smaller as well as larger values, so this exercise again results in a saddle point.

In addition, the system designer would be interested in keeping the sizes of the RF switches for T3 and T4 trunking small. In particular, one would like to minimize the total number of control lines  $m_3^2 + m_4^2$  (see appendix A) for groups 3 and 4. Hence, with eqs. (B.2) and (B.4)

$$\begin{aligned} 0 &= \frac{d}{dM_3} \sum_{i=3}^4 M_i^2 N_i^2 = \frac{d}{dM_3} (M_3^2 N_3^2 + (M - M_1 - M_2 - M_3)^2 N_4^2) \\ &= 2(M_3 N_3^2 - M_4 N_4^2) \end{aligned}$$

a relation also obtainable from differentiating with respect to  $M_4$ , by symmetry. Since the second derivative is non-negative, this implies that the number of FDMA carriers in groups 3 and 4 should be chosen so that

$$N_4/N_3 = (M_3/M_4)^{1/2} \quad (B.16)$$

in order to minimize the complexity of the RF switches. This same approach can be applied to groups 1 and 2 without crosstrunking to yield

$$N_2/N_1 = (M_1/M_2)^{1/2}. \quad (B.17)$$

Finally, one might like to minimize the  $q$  of eq. (B.7), taken as a measure of satellite complexity, subject to the constraints of eqs. (B.2) and (B.3). With eqs. (B.8), (B.11) and (B.12), the Lagrange multiplier technique again leads to saddle points with

$$T_1 = T_2 \text{ and } T_3 = T_4 \quad (B.18)$$

i.e., equal length bursts (slots or dwell times) for groups 1 and 2 and for groups 3 and 4. Any convenient burst length yielding reasonable frame times according to eq. (B.1) is appropriate.

For this model, one should select the numbers of beams and bandwidths to be heavily weighted in favor of a single group with the maximum bandwidth efficiency in order to increase throughput. A true maximum is impossible unless one is interested in serving only one class of user. Reasonable throughput for interesting user representations can result with eqs. (B.10) and (B.12), however. Digital and RF switch complexity can be minimized by satisfying eqs. (B.17) and (B.16), respectively, as nearly as possible once the numbers of beams are determined.

## APPENDIX C

### MODULATION CHARACTERISTICS

Statistics on the bandwidth and detection efficiencies of several important modulations of interest in satellite communications are provided in this appendix. In particular, the performance of various forms of M-ary phase shift keying (MPSK), M-ary frequency shift keying (MFSK), and continuous phase frequency shift keying (CPFSK) will be compared to Shannon's theoretical limit in additive white Gaussian noise (AWGN).

#### SHANNON LIMIT

Let the bandwidth efficiency factor  $\beta$  be defined as the ratio of channel capacity  $C$  in bps to channel bandwidth  $W$  in Hz. Then Shannon's classic formula  $C = W \log_2 (1 + E_b C / N_0 W)$  for AWGN can be rewritten as

$$\beta = \log_2 (1 + \beta E_b / N_0) \quad (C.1)$$

where  $R=C$  is the information rate in bps,  $E_b$  is the received energy per information bit, and  $N_0$  is the single-sided noise power spectral density. This represents the existence of a communication scheme with an arbitrarily small probability of bit error ( $P_b$ ) at the highest possible information rate.

As Bedrosian points out [Bedrosian, 1977], the energy per bit is usually taken as the received energy per data bit  $E'_b$  in bit error rate curves of  $P_b$  versus energy to noise density ratio for practical modulation schemes. The true  $E_b$  is related to the usual  $E'_b$  as  $E_b = E'_b R'/R$ , where  $R'$  is the data rate in b/s. For sufficiently small error probability, the two rates and the two energies are approximately equal, e.g., for the binary symmetric channel with cross-over probability  $P_b = 0.001$ ,

$$\begin{aligned} R/R' &= 1 + P_b \log_2 P_b + (1 - P_b) \log_2 (1 - P_b) \\ &\approx 0.989. \end{aligned}$$

Therefore,  $E_b/N_0$  and  $R$  are within about 1% or 0.05 dB of  $E'_b/N_0$  and  $R'$ , respectively, for  $P_b \lesssim 0.001$ .

A graph of  $\beta$  versus  $E_b/N_0$  derived from eq. (C.1) is used as a measure of excellence in the bandwidth-power tradeoffs for various modulation schemes.

#### M-ARY PHASE AND FREQUENCY SHIFT KEYING

For the usual performance curves for M elementary digital signals [Cahn, 1959; Wozencraft, Jacobs, 1965], it is convenient to convert the symbol error probability  $P_s$  and received signal energy  $E_s$  to the equivalent  $P_b$  and  $E_b$  for sake of comparison with different alphabet sizes M and other modulation schemes. If M is restricted to a power of two, then by definition

$$E_b = E_s / \log_2 M \quad (C.2)$$

because  $\log_2 M$  bits per symbol are transmitted. The conversion between  $P_s$  and  $P_b$  is less obvious.

Assume that the bits comprising M equally likely symbols are Gray coded so that adjacent symbols have  $(\log_2 M) - 1$  bits in common and one bit in disagreement. Such a coding is always possible. Then for sufficiently small  $P_s$ , one can write

$$P_b \approx P_s / \log_2 M. \quad (C.3)$$

This follows because a symbol error essentially occurs only if a symbol adjacent to the one transmitted is received; only one of  $\log_2 M$  bits can be in error.

Thus, curves of  $P_s$  versus  $E_s/N_0$  for constant M can be converted to curves of  $P_b$  versus  $E_b/N_0$  using eqs. (C.2) and (C.3) provided M is a power of two and  $P_s$  is sufficiently small. This works for both MPSK and MFSK. The bandwidth efficiency of these modulations remains to be determined.

Let  $T_s$  be the duration of an elementary signal. With MPSK, the bandwidth is defined as  $1/T_s$ , or precisely half the width of the double-sided main spectral lobe, i.e., the single-sided zero-crossing bandwidth. Since the data rate is

$$R = (\log_2 M) / T_s \text{ bits/sec} \quad (C.4)$$

the bandwidth efficiency becomes simply

$$\beta = R / (1/T_s) = \log_2 M \text{ (MPSK)}. \quad (C.5)$$

The  $M=2$  signals for binary PSK (BPSK) are antipodal in the sense that they are negatives of one another. With a perfect absolute phase (as well as carrier frequency and symbol timing) reference at the (coherent) receiver, the BPSK error probability in AWGN is

$$P_b = \text{erfc}((E_b/N_0)^{1/2})/2 \text{ (coherent BPSK or QPSK).} \quad (\text{C.6})$$

If the  $M=4$  signals of quadriphase shift keying (QPSK) are detected coherently in phase quadrature, the bit error probability of eq. (C.6) is obtained exactly. Larger  $E_b/N_0$ 's are required to achieve the same MPSK error probability with higher order alphabets.

The elementary signals of duration  $T_s$  for MFSK are mutually orthogonal if the  $M$  frequencies are spaced by  $1/T_s$  (Hz) and no phase constraints are imposed. In this case the performance of a coherent binary receiver is 3 dB inferior to that for antipodal signals

$$P_b = \text{erfc}((E_b/2N_0)^{1/2})/2 \text{ (coherent BFSK).} \quad (\text{C.7})$$

However, in contrast to MPSK, the required  $E_b/N_0$  for MFSK decreases with increasing  $M$  for a fixed  $P_b$ .

In this appendix, the bandwidth of MFSK modulation is defined as  $(M-1)/T_s$ , based on the frequency spacing of the orthogonal signals. Note that with the same data rate, the bandwidths of BPSK and BFSK are identical under these definitions. Hence, the bandwidth efficiency of MFSK is

$$\beta = R/((M-1)/T_s) = (\log_2 M)/(M-1) \text{ (MFSK)} \quad (\text{C.8})$$

which is less than that for MPSK for  $M$  larger than two.

Detection efficiency will be degraded without a perfect absolute phase reference at the receiver. If the rate of change in received carrier phase reference is sufficiently slow, MPSK can be demodulated non-coherently by phase comparison techniques. The information is differentially encoded and decoded as phase changes between successive symbols at the transmitter and receiver, respectively, to avoid symbol ambiguities. There is roughly a 1 dB degradation in  $E_b/N_0$  for the typical binary case and an asymptotic 2.3 dB loss for  $M=4$ . Non-coherent (no phase information) BFSK performance is 3 dB worse than phase comparison BPSK.

## MSK AND OFFSET QPSK

Minimum shift keying (MSK) and offset or staggered QPSK (SQPSK) modulation waveforms are shown in figure C-1, along with QPSK. Here  $T = 1/R$ , the reciprocal of the data rate. Note that SQPSK is the same as QPSK except that one of the quadrature data streams has been shifted by  $T$ . MSK is the same as SQPSK except that a sinusoidal elementary signal shape rather than a rectangle is employed. Each of these modulations has a constant envelope but is composed of antipodal elementary signals like BPSK if coherently detected on independent quadrature channels. Thus, the performance of eq. (C.6) can be achieved.

Note that MSK has continuous phase. The transmitted waveform can also be described as continuous phase binary FSK with a deviation ratio  $h = \Delta f T = 1/2$ , where  $\Delta f$  is the separation between the two signaling frequencies. Continuous (frequency) shift keying (CSK) denotes any continuous frequency modulation. An example of CSK is sinusoidal frequency shift keying (SFSK), where the sinusoidal phase shape for MSK of figure C-1 is replaced by one with two continuous derivatives. The total signal still remains one of constant envelope and sinusoidal shaped frequency,  $\omega(t) = \dot{\phi}(t) = \pi(1 - \cos(2\pi t/T))/2T$ , for  $t$  between  $-T$  and  $T$ , for example.

Although SQPSK has the same spectrum as QPSK, the single-sided zero-crossing bandwidth of MSK is  $3R/4$ . With this definition of bandwidth,

$$\beta = R/(3R/4) = 4/3 \text{ (MSK)}. \quad (C.9)$$

Keep in mind, however, that because of the orders of continuity the spectral sidelobes of MSK and SFSK roll-off as  $\omega^{-4}$  and  $\omega^{-8}$ , respectively, while those for QPSK and SQPSK vary as only  $\omega^{-2}$ . Therefore, MSK and SFSK tend to yield much lower crosstalk in FDMA situations.

## OTHER MODULATIONS

So far, all modulations described have been detected on a symbol basis with no memory. Continuous phase FSK (CPFSK) modulation can be improved with additional receiver complexity by introducing memory. That is, a decision on a current received symbol is based on several, e.g.,  $n-1$ , past decisions. One achieves the performance of an error correcting code, without the usual bandwidth expansion for redundancy, by taking advantage of the continuous phase constraints on the modulation. Schonhoff has

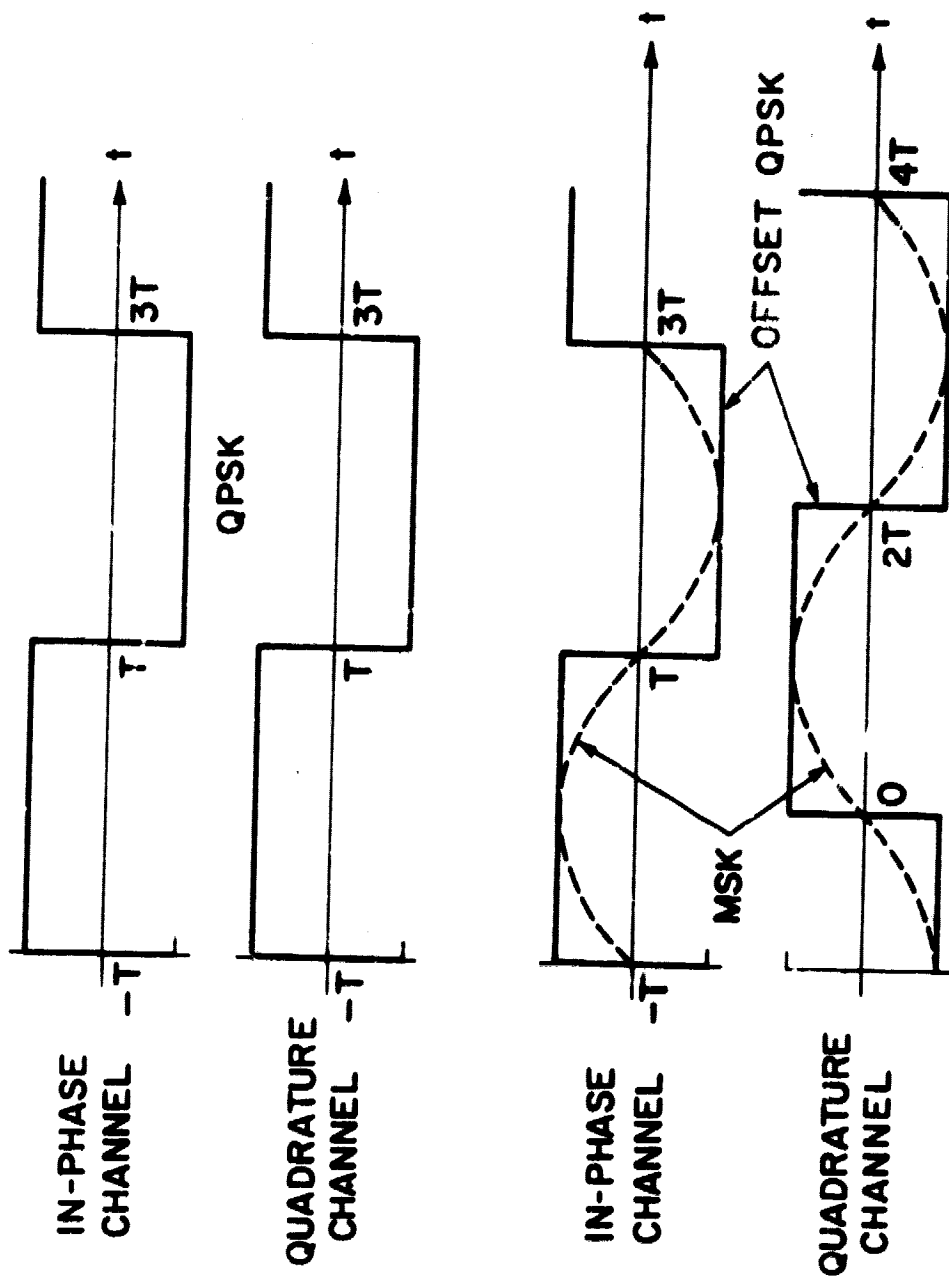


Figure C-1. Baseband Modulation Waveforms

achieved  $\beta$ 's of about 3 at  $E_b/N_0$ 's of approximately 8 dB for  $P_b = 0.00001$  and CPFSK with  $n=3$ , under a 13 dB bandwidth definition [Schonhoff, 1975]. This appears to yield modulations impressively close (within about 5 dB) of the power-bandwidth limit of Shannon.

The bandwidth and detection efficiencies of all the modulations discussed are plotted in figure C-2 for a single channel in AWGN. Digital equivalents of continuous (analog) FM and AM/FM hybrid or QAM schemes with the indicated numbers of quantization levels are also shown [Cuccia, 1978]. Although the QAM modulations have better bandwidth efficiency than MPSK with  $M > 8$ , some form of automatic gain control (AGC) is required at the receiver. This would add to the complexity of a satellite system. Observe that M-ary orthogonal signaling embodied by MFSK is not bandwidth efficient.

The memoryless systems QPSK and SQPSK are quite respectable performers. MSK is also good. The rather arbitrary bandwidth definition used here rather unfairly penalizes MSK. It is well known that MSK usually performs at least as well as SQPSK on a single channel basis; the signal bandwidth does not tell the whole story.



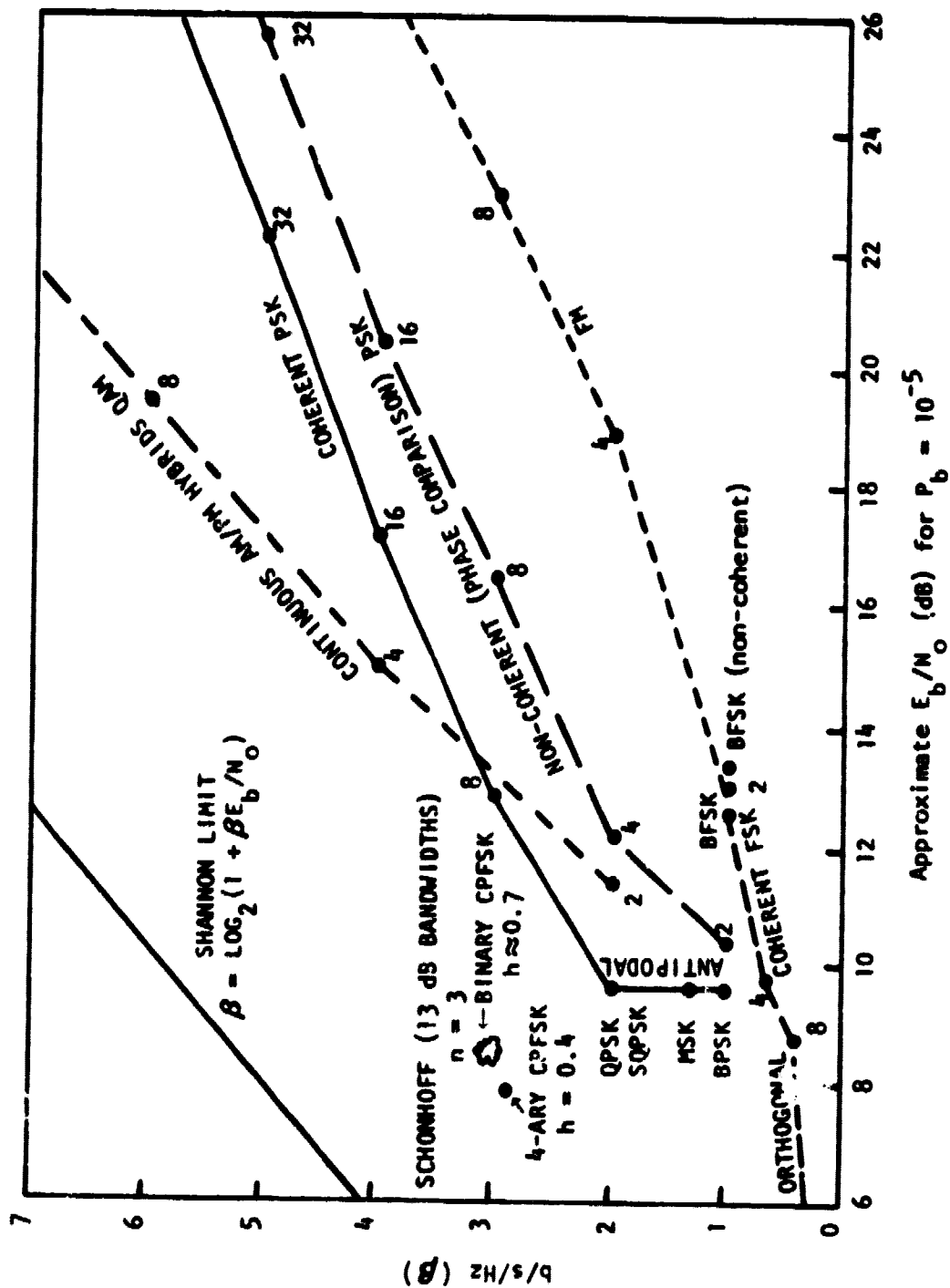


Figure C-2. Single Channel Bandwidth-Power Tradeoffs

#### REFERENCES

Bedrosian, E., "Spectrum Conservation by Efficient Channel Utilization," IEEE Commun. Soc. Magazine, Vol. 15, March 1977, pp. 20-27.

Cahn, C. R., "Performance of Digital Phase-Modulation Communications Systems, Trans. IRE, Vol. CS-7, May 1959, pp. 3-6.

Cuccia, C. L., "Bandwidth Conservation is Essential," Microwave Systems News, Vol. 3, October 1978, pp. 67-72, 161.

Schonhoff, T. A., "Bandwidth vs. Performance Considerations for CPFSK," Proc. Nat. Telecommun. Conf., New Orleans, 1-3 December 1975, pp. 38-1 through 38-5.

Wozencraft, J. M., Jacobs, I. M., Principles of Communications Engineering, Wiley, New York, 1965, pp. 248-252.

## APPENDIX D

### DISCUSSION OF $R_0$ AS A CHANNEL PERFORMANCE MEASURE

A significant result of information theory has been the determination of bounds on the attainable probability of error for a given transmission channel as a function of data rate, for rates less than channel capacity. These bounds are important because they provide relatively simple, but sharp, estimates of average error rates which in general cannot be computed precisely. [Fano, 1963] and [Gallager, 1968] have shown that for a broad class of channels the minimum attainable error probability,  $P[\xi]$ , is bounded above and below by the expression

$$A_1 2^{-NE_L(R)} \leq P[\xi] \leq A_2 2^{-NE(R)} \quad (D.1)$$

in which  $N$  is related to the constraint length of the code and  $R$  is the information rate in bits per channel symbol. In principle, an arbitrary small error rate can be obtained by choosing  $N$  sufficiently large, provided that  $E(R) > 0$ . The coefficient  $A_1$  in eq. (D.1) decreases so slowly as  $N$  is increased that the attainable performance when  $N$  is large is determined essentially by the functions  $E_L(R)$  and  $E(R)$ , called the reliability functions.

The expressions for  $E_L(R)$  and  $E(R)$  depend primarily on the nature of the channel and specifically are positive functions of the channel transition probabilities. For a range of rates,  $R$ , near channel capacity  $E_L(R) = E(R)$ . The function  $E(R)$ , especially in this range, yields a fundamental characterization of a channel and in turn brings out clearly the relationship between error probability, data rate, code complexity, and channel behavior. The larger  $E(R)$ , the better the channel.

In our application we are considering an end-to-end channel from uplink terminal through a satellite to the downlink terminal. This channel is assumed to contain different on-board processors as part of the channel with the behavior of the end-to-end channel strongly dependent on the nature of the processor. The  $E(R)$  function provides a theoretical framework within which to discuss intelligently the relative merits of each processing technique; however, because of the complexity of the processing channel we will

use the function  $R_0$  which is related to  $E(R)$  but computationally easier to use. The relationship between  $E(R)$  and  $R_0$  is now shown.

A good derivation of the bound given in eq. (D.1) is given by [Gallager, 1965]; the exponent  $E(R)$  is defined as:

$$E(R) = \max_{0 \leq \rho \leq 1} \max_P [E_0(\rho, P) - \rho R] \quad (D.2)$$

where  $\rho$  is an arbitrary number,  $0 \leq \rho \leq 1$  and  $P = (p_1, p_2, \dots, p_M)$  is a probability vector of the input alphabet. The function  $E_0$  is an explicit function of the channel and for the discrete memoryless channel (DMC) is given by:

$$E_0(\rho, P) = -\log_2 \sum_{j=1}^J \left( \sum_{k=1}^M p_k q_{jk}^{1/(1+\rho)} \right)^{1+\rho}$$

where  $q_{jk}$  is the transition probability between the channel input and output. If we look carefully at eq. (D.2) we see that  $E_0(\rho, P) - \rho R$  is just a linear function of  $R$  with slope  $-\rho$  and intercept  $E_0(\rho, P)$  for fixed  $\rho$ . Thus,  $E(R)$  is simply the upper envelope of a set of straight lines as shown in figure D-1 and therefore represents the exponent of the tightest bound in eq. (D.1).

The procedure for performing the maximization indicated in eq. (D.2) to obtain the  $E(R)$  function is given by [Gallager, 1965]. However, we see in figure D-1 the  $E(R)$  function can be approximated by a straight line function by fixed  $\rho$  and in particular, by setting  $\rho = 1$  we obtain a function with special properties. That is,

$$E(R) \sim E'(R) = \max_P [E_0(1, P) - R]$$

where

$$E_0(1, P) = R_0(P)$$

and for the DMC,

$$R_0(P) = \max_P \left[ -\log_2 \sum_{j=1}^J \left( \sum_{k=1}^M p_k q_{jk}^{1/2} \right)^2 \right]$$

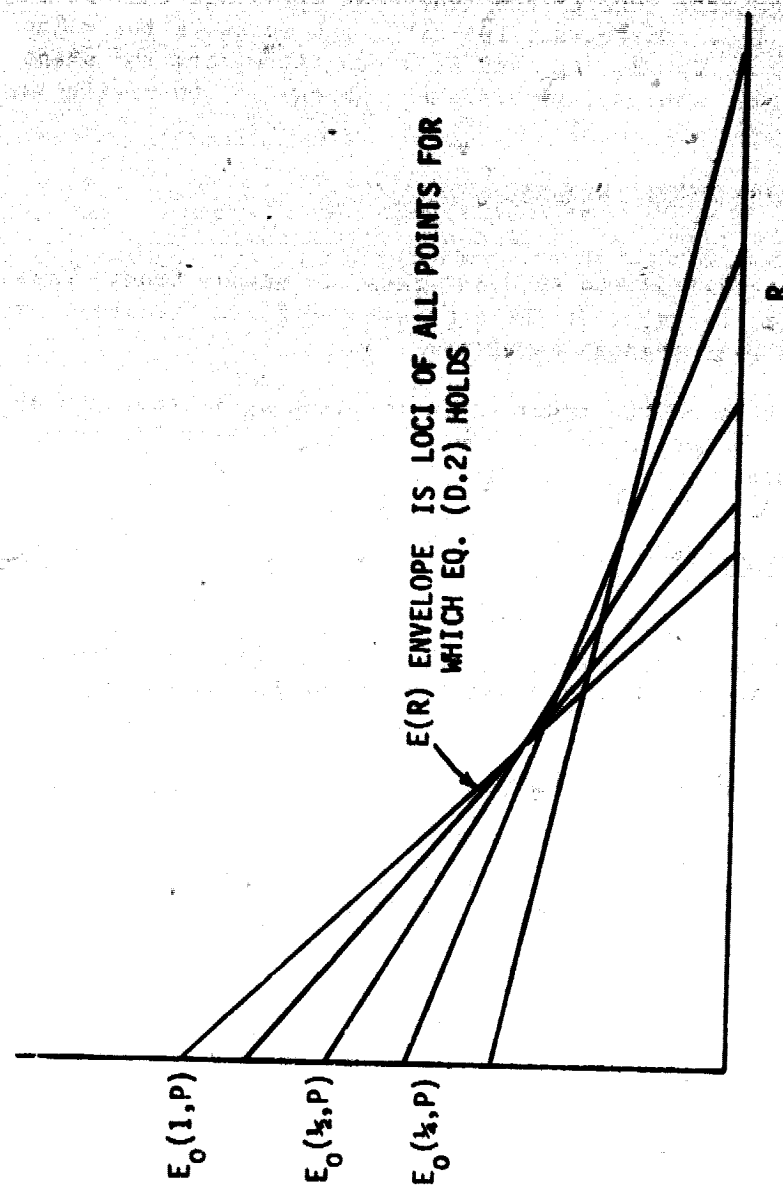


Figure D-1. Exponent Rate Curves as Envelope of Straight Lines

Generally the maximization of  $E_0$  over  $P$  is not done and  $P$  is assumed  $1/M$ . For the binary case ( $M=2$ ) the uniform distribution always maximizes  $E_0$  and this is also true for other cases of practical interest where the modulation signal set and the demodulator decision regions are reasonably symmetric.

The generic form of the functions  $E_L(R)$  and  $E(R)$  is illustrated in figure D-2. Of course, for different channels the values of the parameters  $C$ ,  $R_0$ ,  $R_C$ ,  $R_e$ , and  $E(0)$  are different; the shape of the curves shown, however, is extremely general.\* In particular, the equality

$$E_L(R) = E(R); \quad R_C \leq R \leq C \quad (D.3)$$

where  $R_C$  is called the critical rate, is always true. Thus, the exponential behavior of the attainable  $P[\xi]$  is precisely determined for rates near channel capacity.

The form of the error bound in which  $E_0$  is replaced by  $R_0$  is called the union bound and developed in [Wozencraft, 1965]. The union bound,

$$P[\xi] \leq 2^{-N[R_0 - R]} \quad (D.4)$$

is indicated in figure D-2 by the dashed line. The exponent  $E(R)$  always coincides with the union exponent for  $R_e \leq R \leq R_C$ , but in general,

$$E(R) > R_0 - R; \quad 0 \leq R \leq R_e \quad (D.5)$$

In addition, [Viterbi, 1967] has shown that for convolutional codes and maximum likelihood decoding,

$$E(R) = \begin{cases} R_0 & ; 0 \leq R < R_0 \\ E_0(\rho, P) & ; R_0 \leq R < C \end{cases}$$

---

\*In certain instances the curves may exhibit degeneracies. As an example, for the infinite bandwidth white Gaussian noise channel  $E(0) = R_0$  and  $E_L(R) = E(R)$  for all  $R$ .

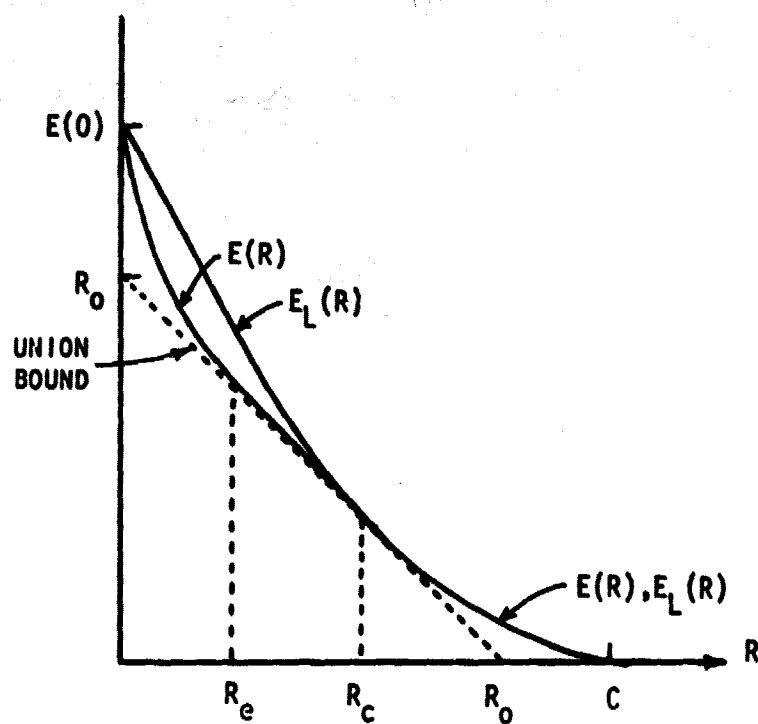


Figure D-2. Typical Channel Reliability Functions

The curves  $E_L(R)$  and  $E(R)$  for a specific channel embody detailed knowledge of the attainable error performance. Although less detailed, the knowledge conveyed just by the value of  $R_0$  is also exceedingly informative. In particular, figure D-2 illustrates that the union bound of eq. (D.4) is exponentially equivalent to the lower bound of eq. (D.1) for  $R \approx R_0$ . Thus, the value of  $R_0$  provides an accurate characterization of the exponential error behavior attainable at rates  $R_0$ .

The advantage in simplicity to be gained from using a single-parameter descriptor is obvious, and in our study of the implementation of on-board processing for coded data we focus attention primarily on  $R_0$ .

The channel we are interested in consists of  $M$  discrete inputs and  $M$  continuous outputs which may be represented by a vector  $y = (y_1 \dots y_M)$ ;  $R_0$  for such a channel is given by

$$R_0 = \max_P \left\{ -\log_2 \int_{-\infty}^{\infty} dy \left[ \sum_{k=1}^M p_k \sqrt{q(y/k)} \right]^2 \right\} \quad (D.6)$$

where  $p_k$ ,  $1 \leq k \leq M$  is the probability of transmitting the  $k$ th orthogonal signal and  $q(y/k)$  is the conditional probability of the output if the  $k$ th signal was transmitted.

Under the assumption that the outputs conditioned on a particular input are statistically independent, the conditional probability factors into a product of terms:

$$q(y/k) = p_{s+n}(y_k) \prod_{\substack{j=1 \\ j \neq k}}^M p_n(y_j) \quad (D.7)$$

where  $p_{s+n}(y_k)$  is the probability density of the output corresponding to the transmitted signal and  $p_n(y_j)$  corresponds to the remaining outputs. From this it can be shown that:

$$R_0 = \log_2 \frac{M}{1 + (M-1) \exp \left( -\frac{E}{2N_0} \right)}$$



The parameter  $R_0$  is used to compare the system model with on-board processing and a linear repeater satellite to determine the comparative performance of the various processing schemes.

#### REFERENCES

Fano, R. M., "A Heuristic Discussion of Probabilistic Decoding,"  
IEEE Trans. on Info. Th., Vol. IT-9, pp. 64-74, April 1963.

Gallager, R. G., Information Theory and Reliable Communications, New  
York: Wiley, 1968.

Gallager, R. G., "A Simple Derivation of the Coding Theorem and Some  
Application," IEEE Trans. on Info. Th., Vol. IT-11, pp. 3-18,  
January 1965.

Viterbi, A. J., "Error Bounds for Convolutional Codes and an  
Asymptotically Optimum Decoding Algorithm," IEEE Trans. on Info.  
Th., Vol. IT-13, No. 2, April 1976, pp. 260-269.

Wozencraft, J. M., and Jacobs, I. H., Principles of Communication  
Engineering, New York: Wiley, 1965.

## GLOSSARY

ADCCP	advanced data communication control procedures
AGC	automatic gain control
AM	amplitude modulation
ANSI	American National Standards Institute
ARQ	automatic repeat request
AWGN	additive white Gaussian noise
B.O.	back off
BC	billing center
BER	bit error rate
BOM	beginning of mission
BPSK	binary phase shift keyed
CCD	charge-coupled device
CCIS	Common Channel Interoffice Signaling
CCITT	International Telephone and Telegraph Consultive Committee
CDC	Control Data Corporation
CDCP	control data communications control procedure
CDMA	code division multiple access
CONUS	continental United States
CPFSK	continuous phase frequency shift keyed
CQPSK	coherent quadri-phase shift keyed
CSK	continuous shift keying
CSMA	carrier sense multiple access
DA	demand assignment
DAC	demand assignment controller
DAMA	demand assignment multiple access
DB	data base
DDCMP	digital data communications message protocol
DDD	direct distance dialing
DEP	decoding-encoding processing
DFT	digital Fourier transform
DMA	direct-memory-access
ECL	emitter-coupled logic
EHF	extremely high frequency
EIRP	effective isotropic radiated power
EMI	electromagnetic interference
ENR	energy-to-noise ratio
FDM	frequency division multiplex
FDMA	frequency division multiple access
FET	field effect transistor
FH	frequency hopping
FIFO	first-in first-out
FM	frequency modulation
FSK	frequency shift keyed

G/T	gain-to-noise temperature ratio
GaAs	gallium arsenide
HDLC	high level data link control
HDP	hard decision processing
HPA	high power amplifier
IC	integrated circuit
IF	intermediate frequency
IM	intermodulation
ISI	intersymbol interference
ISO	International Organization for Standards
LeRC	Lewis Research Center
LNA	low noise amplifier
LRPT	linear repeater
LSI	large scale integration
MA	multiple access
MESFET	metal epitaxial semiconductor field effect transistor
MF	multifrequency
MFSK	M-ary frequency shift keying
MOS	metal oxide semiconductor
MPSK	M-ary phase shift keying
MSK	minimum shift keyed
NASA	National Aeronautics and Space Administration
NMOS	N-type MOS
NRE	non-recurring
O&M	operations and maintenance
OBC	on-board controller
P/S	parallel-to-serial
PCB	printed circuit board
PCSFSK	phase comparison sinusoidal frequency shift keying
pfd	power flux density
PM	phase modulation
PN	pseudonoise
QPSK	quadri-phase shift keying
RA	random access
RAM	random access memory
RCC	regional control center
RE	recurring
RF	radio frequency
RFI	radio frequency interference
S/P	serial-to-parallel
SAW	surface acoustic wave
SCC	system control center
SCMC	system control and management center
SCPC	single channel per carrier
SDLC	synchronous data link control
SDMA	space division multiple access
SF	single frequency

<b>SIPO</b>	<b>serial input to parallel output</b>
<b>SMC</b>	<b>system management center</b>
<b>SMSA</b>	<b>Standard Metropolitan Statistical Area</b>
<b>SO</b>	<b>switching office</b>
<b>SPMT</b>	<b>single pole multiple throw</b>
<b>SQPSK</b>	<b>staggered quadri-phase shift keyed</b>
<b>SS-TDMA</b>	<b>satellite switched time division multiple access</b>
<b>SSMA</b>	<b>spread spectrum multiple access</b>
<b>STP</b>	<b>signal transfer point</b>
<b>SWR</b>	<b>standing wave ratio</b>
<b>TDM</b>	<b>time division multiplex</b>
<b>TDMA</b>	<b>time division multiple access</b>
<b>TMC</b>	<b>telemetry monitoring center</b>
<b>TSPS</b>	<b>Traffic Service Position Systems</b>
<b>TTL</b>	<b>transistor-to-transistor logic</b>
<b>TWT</b>	<b>traveling wave tube</b>
<b>TWTA</b>	<b>traveling wave tube amplifier</b>
<b>TX</b>	<b>transmit</b>
<b>VSWR</b>	<b>voltage standing wave ratio</b>